

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 16, 2002, 23:19:15 ; Search time 4188 Seconds  
(without alignments)  
16601.386 Million cell updates/sec

Title: US-08-961-083-55

Perfect score: 2389

Sequence: 1 TTCTTAGCAGTGGAGCTCT.....TAAGTAAGCAAAATAAAC 2389

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_pro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_pro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_huv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_fod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2388	100.0	2389	6	ARI20265	ARI20265 Sequence
2	2388	100.0	2341	1	AF291695	AF291695 Streptococ
3	2388	100.0	8195	6	BD003774	BD003774 Polynucle
4	2388	100.0	10256	1	AE007418	AE007418 Streptococ
5	2365	99.0	10320	1	AE008479	AE008479 Streptococ
6	2328	97.5	20035	2	SPNEU1915	SPNEU1915 Streptococ
7	1434	60.0	2535	1	AF340221	AF340221 Streptococ
8	1374	57.5	2523	6	AX343072	AX343072 Sequence
9	1374	57.5	2647	6	AX343073	AX343073 Sequence
10	991	41.5	2457	1	AF318954	AF318954 Streptococ
11	990	41.4	2517	1	AF318955	AF318955 Streptococ
12	990	41.4	11931	1	AE007403	AE007403 Streptococ
13	988	41.4	232807	2	SPNEU1901	SPNEU1901 Streptococ
14	987	41.3	2290	6	ARI20270	ARI20270 Sequence
15	980	41.1	2639	6	AX343074	AX343074 Sequence
16	962	40.3	12372	1	AE008464	AE008464 Streptococ
17	869	36.4	2166	12	AF340222	AF340222 Synthetic
18	835	35.0	75874	2	SPNEU1907	SPNEU1907 Streptococ
19	674	28.2	702	2	SPNEU1929	SPNEU1929 Streptococ
20	653	27.4	2359	6	BD003923	BD003923 Polynucle
21	385	16.1	1342	6	ARI20328	ARI20328 Sequence
22	385	16.1	3120	1	AF318956	AF318956 Streptococ
23	385	16.1	3120	6	AX343070	AX343070 Sequence
24	385	16.1	5048	6	AX343071	AX343071 Sequence
25	385	16.1	6867	6	BD003872	BD003872 Polynucle
26	381	16.0	973	6	BD004035	BD004035 Polynucle
27	347	14.5	2528	6	AX343078	AX343078 Sequence
28	253	10.6	10312	1	AE007404	AE007404 Streptococ
29	243	10.2	1684	6	BD003938	BD003938 Polynucle
30	214	9.0	2478	6	AX088372	AX088372 Sequence
31	214	9.0	10825	1	AE006623	AE006623 Streptococ
32	213	8.9	2478	1	AB073859	AB073859 Streptococ
33	213	8.9	10618	1	AE010110	AE010110 Streptococ
34	211	8.8	1910	1	SAG290952	AJ290952 Streptococ
35	211	8.8	2469	6	AX088376	AX088376 Sequence
36	211	8.8	3501	1	AF062533	AF062533 Streptococ
37	205	8.6	50354	1	AE014169	AE014169 Streptococ
38	188	7.9	492	12	AF340223	AF340223 Synthetic
39	109	4.6	841	6	AR089275	AR089275 Sequence
40	109	4.6	841	6	AR093475	AR093475 Sequence
41	70	3.0	115758	9	AC104634	AC104634 Homo sapi
42	63	2.7	170522	9	AP002387	AP002387 Homo sapi
43	60	2.5	51939	9	AC099341	AC099341 Homo sapi
44	60	2.5	61052	2	AC117074	AC117074 Dictyoste
45	60	2.5	61052	2	AC123513	AC123513 Dictyoste

#### ALIGNMENTS

RESULT 1  
LOCUS ARI20265 2389 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 55 from patent US 6159469.  
ACCESSION ARI20265  
VERSION ARI20265.1 GI:14103841  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2389)  
AUTHORS Choi,G.H., Kunsch,C.A., Barash,S.C., Dillon,P.J., Dougherty,B.,  
Fannon,M.R. and Rosen,C.A.  
TITLE Streptococcus pneumoniae antigens and vaccines  
JOURNAL Patent: US 6159469-A 55 12-DEC-2000;

```
FEATURES                               Location/Qualifiers
source                                1..2389
                                         /organism="unknown"
BASE COUNT      830 a      461 c      486 g      611 t      1 others
ORIGIN
Query Match      100.0%: Score 2388; DB 6; Length 2389;
Best Local Similarity 100.0%: Pred. No. 0;
Matches 2389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTCTTACGCGTTGGGAGCTATCAAGCTGGAACGGTTAAGSAAATTAATCGTTTCCCA 60
Db      1 TTCTTACGCGTTGGGAGCTATCAAGCTGGAACGGTTAAGSAAATTAATCGTTTCCCA 60
QY      61 TTATAGTGGAAAAACAAGCAGCAGCAAAAAAGGAGATTGTACTCTGATGAGTTAGCAA 120
Db      61 TTATAGTGGAAAAACAAGCAGCAGCAAAAAAGGAGATTGTACTCTGATGAGTTAGCAA 120
QY      121 GCGTGAAGGATTCATGCTGAGCAAAATGCTCATCAAGATTAACAGACCAAGGCTATGTAC 180
Db      121 GCGTGAAGGATTCATGCTGAGCAAAATGCTCATCAAGATTAACAGACCAAGGCTATGTAC 180
QY      181 TTCACATGGCGACCATCATCATTTATACATGTAAGGTTCCCTTATGAGCTATCATCAG 240
Db      181 TTCACATGGCGACCATCATCATTTATACATGTAAGGTTCCCTTATGAGCTATCATCAG 240
QY      241 TGAAGATTACTCATGAAGATCCAAACTATAAGCTAAAGATGAGGATTTGTATATGA 300
Db      241 TGAAGATTACTCATGAAGATCCAAACTATAAGCTAAAGATGAGGATTTGTATATGA 300
QY      301 GGTCAAGGCTGATATGTTTATCAAGGTAGATGAAATTCATATGTTTACCTTAAGATGC 360
Db      301 GGTCAAGGCTGATATGTTTATCAAGGTAGATGAAATTCATATGTTTACCTTAAGATGC 360
QY      361 TGGCCACGGGATTAAGCTCGTACAAAAGAGAAATTCATCCGCAAAAACAAGAGATAG 420
Db      361 TGGCCACGGGATTAAGCTCGTACAAAAGAGAAATTCATCCGCAAAAACAAGAGATAG 420
QY      421 TCACATCTGTGAAGGTGGAACCTCAAGAAACGATGGTGTCTGTGCGTGGCAGCTTCGCA 480
Db      421 TCACATCTGTGAAGGTGGAACCTCAAGAAACGATGGTGTCTGTGCGTGGCAGCTTCGCA 480
QY      481 AGGAGCCTTACTACAGATGATGGTTATATCTTTAATGCTTCTGATATATAGAGATAC 540
Db      481 AGGAGCCTTACTACAGATGATGGTTATATCTTTAATGCTTCTGATATATAGAGATAC 540
QY      541 TGGTATGCTTATATCGTTCCTCATGAGAGATCATTTACATTCCTTAAGATAGATT 600
Db      541 TGGTATGCTTATATCGTTCCTCATGAGAGATCATTTACATTCCTTAAGATAGATT 600
QY      601 ATTCAGCTAGCGAGTGGCTGTGTCAGAAAGCCTTCTATCTGTGTCGAGAAATCTGTCAAA 660
Db      601 ATTCAGCTAGCGAGTGGCTGTGTCAGAAAGCCTTCTATCTGTGTCGAGAAATCTGTCAAA 660
QY      661 TTCAGAAGACCTTATCCGCCGCAAAATAGCGATTAACACTTAAAGAAACAACTGGGTACTTTC 720
Db      661 TTCAGAAGACCTTATCCGCCGCAAAATAGCGATTAACACTTAAAGAAACAACTGGGTACTTTC 720
QY      721 TSTAAGCAATCCAGGAACACTACAATACTTAACACAGCAACAAGCAAGCACTAAACAGTCA 780
Db      721 TSTAAGCAATCCAGGAACACTACAATACTTAACACAGCAACAAGCAAGCACTAAACAGTCA 780
QY      781 AGCAAGTCAAAAGTAATGACATTGATAGTCTCTTGAAACAGCTCTACAAACCTGCTTTGAG 840
Db      781 AGCAAGTCAAAAGTAATGACATTGATAGTCTCTTGAAACAGCTCTACAAACCTGCTTTGAG 840
QY      841 TCAAGCAGCATGTAGAAATCTGATGGCTTGTCTTTGATCCAGACAAATATACAAGTGAAC 900
Db      841 TCAAGCAGCATGTAGAAATCTGATGGCTTGTCTTTGATCCAGACAAATATACAAGTGAAC 900
QY      901 AACTAGAGGTGTGAGAGTCCACAGAGAGATCATTTACCACTTCATCCCTTACTCTCAAT 960
Db      901 AACTAGAGGTGTGAGAGTCCACAGAGAGATCATTTACCACTTCATCCCTTACTCTCAAT 960
```

```
QY      961 GTCGTGAATTTGAAGAAGCAATGCTCGTATATATTCCTCCCTTCGTTATCGTTCAACCATTTG 1020
Db      961 GTCGTGAATTTGAAGAAGCAATGCTCGTATATATTCCTCCCTTCGTTATCGTTCAACCATTTG 1020
QY      1021 GGTACCAATTTCAAGGCCAGAAACAACCAAGTCCACACAGCGATCCGGAACTAGTCCAG 1080
Db      1021 GGTACCAATTTCAAGGCCAGAAACAACCAAGTCCACACAGCGATCCGGAACTAGTCCAG 1080
QY      1081 CCCGCACCTGCAACCAATCTTAAATAGACTCAAAATCTTCTTGTTAGTACAGTGGT 1140
Db      1081 CCCGCACCTGCAACCAATCTTAAATAGACTCAAAATCTTCTTGTTAGTACAGTGGT 1140
QY      1141 ACGAAAAGTTGGGGAAGATATGTATTCGAGAAAAAGGCAATCTCTGTTATGCTTTTGC 1200
Db      1141 ACGAAAAGTTGGGGAAGATATGTATTCGAGAAAAAGGCAATCTCTGTTATGCTTTTGC 1200
QY      1201 GAAAGATTACCATCTGAACCTGTTAAAAATCTGTGAACCAAGTTATCCAAAACAAGAGAG 1260
Db      1201 GAAAGATTACCATCTGAACCTGTTAAAAATCTGTGAACCAAGTTATCCAAAACAAGAGAG 1260
QY      1261 TGTTCACACACTTTAAGCTTAAAAAAGAAAATGTTGCTCCCTGAGACCAGAATTTTA 1320
Db      1261 TGTTCACACACTTTAAGCTTAAAAAAGAAAATGTTGCTCCCTGAGACCAGAATTTTA 1320
QY      1321 TGATTAAGCATATATCTGTATTAAGGCTCATAAAGCCTTGTGTTGNAATPAAGGTTGC 1380
Db      1321 TGATTAAGCATATATCTGTATTAAGGCTCATAAAGCCTTGTGTTGNAATPAAGGTTGC 1380
QY      1381 TAATTCGTATTTCCAAAGCCTTAGACAAATTTATAGAACCTTGAATGATTCGACTTA 1440
Db      1381 TAATTCGTATTTCCAAAGCCTTAGACAAATTTATAGAACCTTGAATGATTCGACTTA 1440
QY      1441 TAAAGAAAAATTGTAGATGATTTATGSCATTTCCAGACCAATTAACCATCCAGAGCG 1500
Db      1441 TAAAGAAAAATTGTAGATGATTTATGSCATTTCCAGACCAATTAACCATCCAGAGCG 1500
QY      1501 ACTTGGCAAAACCAAAATTCCTCAAAATGAGTATCTGAAGACGAATTCGTATTTCTCAATT 1560
Db      1501 ACTTGGCAAAACCAAAATTCCTCAAAATGAGTATCTGAAGACGAATTCGTATTTCTCAATT 1560
QY      1561 AGCTGATAGTATACAGCTGATGCTGATCAATTTTGTATGAAACATGATATATATGCTTA 1620
Db      1561 AGCTGATAGTATACAGCTGATGCTGATCAATTTTGTATGAAACATGATATATATGCTTA 1620
QY      1621 TGAAGGAGATGATATGTATTAACGCTCATATGSGGCATAGTCACTGGATTGGAAGAATAG 1680
Db      1621 TGAAGGAGATGATATGTATTAACGCTCATATGSGGCATAGTCACTGGATTGGAAGAATAG 1680
QY      1681 CCTTCTGATTAAGGAAAAAGTTGCAAGCTCAAGCTTACTTAAAGAAAAAGTATCTTACC 1740
Db      1681 CCTTCTGATTAAGGAAAAAGTTGCAAGCTCAAGCTTACTTAAAGAAAAAGTATCTTACC 1740
QY      1741 TCCATCTCCAGACGCGAGATGTTTAAAGCAAAATCCAACTGAGATAGTGCAGCGATTTTA 1800
Db      1741 TCCATCTCCAGACGCGAGATGTTTAAAGCAAAATCCAACTGAGATAGTGCAGCGATTTTA 1800
QY      1801 CAATCGTGTGAAGGGAAGGAAAAAGAAATTCACCTGTTGACCTTCATATATGTTGAGCA 1860
Db      1801 CAATCGTGTGAAGGGAAGGAAAAAGAAATTCACCTGTTGACCTTCATATATGTTGAGCA 1860
QY      1861 TACAGTTGAGGTTAAAAACGGTAAATTTGATTTTCTCTCAATAAGATCATTTACCATATAT 1920
Db      1861 TACAGTTGAGGTTAAAAACGGTAAATTTGATTTTCTCTCAATAAGATCATTTACCATATAT 1920
QY      1921 TAAATTTGCTGTTGATGATCACACATACAAAGTCCAAATGGCTATACCTTGGAGAGA 1980
Db      1921 TAAATTTGCTGTTGATGATCACACATACAAAGTCCAAATGGCTATACCTTGGAGAGA 1980
QY      1981 TTTGTTTGGAGATTAAATACGTATAGTAGAACAACCTTGACGAACGTCACATTTCTAATGA 2040
Db      1981 TTTGTTTGGAGATTAAATACGTATAGTAGAACAACCTTGACGAACGTCACATTTCTAATGA 2040
```

QY 2041 TCGATGGGGCAATGCCAGTGCATGTGTTAGGCAAGAAACACAGTGAATCCAAA 2100  
|||||  
Db 2041 TCGATGGGGCAATGCCAGTGCATGTGTTAGGCAAGAAACACAGTGAATCCAAA 2100  
QY 2101 TAAGAACTTCAAGCGGATGGAAGGCCAGTGAAGAAACACCTCTGAGCCAGAATCCC 2160  
|||||  
Db 2101 TAAAGAACTTCAAGCGGATGGAAGGCCAGTGAAGAAACACCTCTGAGCCAGAATCCC 2160  
QY 2161 TCAAGTGAAGAGCTGAAAAGAGAGAGCCCACTCAAGAAAGCAGAGTTTGTGGGAA 2220  
|||||  
Db 2161 TCAAGTGAAGAGCTGAAAAGAGAGAGCCCACTCAAGAAAGCAGAGTTTGTGGGAA 2220  
QY 2221 AGTAAGCAATTCATGCTGTAAGCCCAATGCAACAAACTCTAGCTGTTAGCAAAATA 2280  
|||||  
Db 2221 AGTAAGCAATTCATGCTGTAAGCCCAATGCAACAAACTCTAGCTGTTAGCAAAATA 2280  
QY 2281 TTTCACCTCTCAATTTATGATACATATGATATGTCGAGAGACAGAAATAATTTCTGC 2340  
|||||  
Db 2281 TTTCACCTCTCAATTTATGATACATATGATATGTCGAGAGACAGAAATAATTTCTGC 2340  
QY 2341 GTTGTAAAAGAGATATCTCATCTGTATAGTAAAGAAAATAATTAAC 2389  
|||||  
Db 2341 GTTGTAAAAGAGATATCTCATCTGTATAGTAAAGAAAATAATTAAC 2389  
RESULT 2  
AF291695 2541 bp DNA linear BCT 15-MAR-2001  
LOCUS  
DEFINITION Streptococcus pneumoniae pneumococcal histidine triad A protein  
ACCESSION AF291695  
VERSION AF291695.1 GI:13345012  
KEYWORDS  
SOURCE Streptococcus pneumoniae.  
ORGANISM Streptococcus pneumoniae.  
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
Streptococcus.  
REFERENCE 1 (bases 1 to 2541)  
Witzmann, T. M., Heinrichs, J. H., Adamou, J. E., Erwin, A. L., Kunsch, C.,  
Choi, G. H., Barash, S. C., Rosen, C. A., Masure, H. R., Tomanen, E.,  
Gayle, A., Brehm, Y. A., Walsh, W., Barren, P., Lathigra, R., Hanson, M.,  
Langemann, S., Johnson, S. and Koenig, S.  
Use of a whole genome approach to identify vaccine molecules  
affording protection against Streptococcus pneumoniae infection  
Infect. Immun. 69 (3), 1593-1598 (2001)  
TITLE  
JOURNAL MEDLINE  
PUBMED 2116976  
2 (bases 1 to 2541)  
REFERENCE Choi, G. H.  
AUTHORS Direct Submission  
TITLE Submitted (01-AUG-2000) Molecular Biology, Human Genome Sciences,  
JOURNAL Inc., 9410 Key West Ave., Rockville, MD 20850, USA  
FEATURES  
source  
1. .2541  
/organism="Streptococcus pneumoniae"  
/strain="N4"  
/db\_xref="taxon:1313"  
91. .2541  
/gene="phtA"  
91. .2541  
/gene="phtA"  
/note="phtA"  
/codon\_start=1  
/transl\_table=11  
/product="pneumococcal histidine triad A protein"  
/protein\_id="AAK1915.1"  
/db\_xref="GI:13345013"  
/translation="MKINKKYLIVGSAALILSVSELYGQARTKENVKNSYIDK  
QATOKTENLTPDEVSKREGINAOIYIKITIDGQYVSHGHYHYKQVYDAIISSE  
LTKMDPNYKIKEDDIYEVKGYIVKDGKTYIYKDAHADVVRKEELNKRKOEHS  
QHEGCTPRNDGAVVALARSOGRTITDGTIFNLSIDIETDGDVIVPHGDHYIPEK  
ELISASLEAAAEAFLSGRNLSNRTYRKQSDMTSPNWPVSVNGTPTNTNSN  
TNSOASOSNDIDSLIKOLYKPLPSORHVESDGLVDPAQITSRTAGVAVPHGDHYF

IPYSONSELEERIIARIIPLYRSNHNWPDSPRPOSPPEPSPGPOPANLKDIDSN  
SLAVDQEVKGVGEVEEKGIRYFAKLPSEYKNLESKJESKQESVHTTAKKE  
NSAPROEDYRKAYNLTTEHAKRLPFYKNKNSPFOALDKLELNESTKRELVNDL  
LAFIAPITHEBERIGKRSQIEYTEDEYRIQLDKTTSISGTIFDEHDIISDGDYV  
TPHGHSHWIGKDSLDERKVAQATYKREGILIPSPDADAVKANPFGDASAALYNKVK  
GKRIPLVRLPYVVEHTEYKNGNLIPPHDHNHNTKFAFEDHTYKAPNGYLEDLF  
ATIKYVEVHDEPHERSHNDGNSHEVLKPSDEPNKFKADEPVEETPAPEVP  
OVEDKEVEAOLKEAEVYLAKVTDSSLKANETELAGLRNLTQINDNNSIMAEKXL  
LALLKSNPSVSKEKIN"  
BASE COUNT 888 a 476 c 516 g 660 t 1 others  
ORIGIN  
Query Match 100.0%; Score 2388; DB 1; Length 2541;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTCTTACGAGTTGGGACTGTATCAAGCTAGACGCGTTAAGAAAATATCGTGTTCCTA 60  
|||||  
Db 150 TTCTTACGAGTTGGGACTGTATCAAGCTAGACGCGTTAAGAAAATATCGTGTTCCTA 209  
QY 61 TATAGTGAAGAAACAGAGCGCAAAAAGGAGAAATTTGACTCTGATGAGGTTAGCAA 120  
|||||  
Db 210 TATAGTGAAGAAACAGAGCGCAAAAAGGAGAAATTTGACTCTGATGAGGTTAGCAA 269  
QY 121 GCGTGAAGAAATCAATGCTGAGCAAAATCGTCATCAAGATAACAGACCAAGCTATGTCAC 180  
|||||  
Db 270 GCGTGAAGAAATCAATGCTGAGCAAAATCGTCATCAAGATAACAGACCAAGCTATGTCAC 329  
QY 181 TTACATGGCGACACATATCATTTATCAATGTAAGGTCCTTATGACGCTATCATCAG 240  
|||||  
Db 330 TTACATGGCGACACATATCATTTATCAATGTAAGGTCCTTATGACGCTATCATCAG 389  
QY 241 TGAAGATTTACTCATGAAAGATCCAAACTATACCTAAAGATAGATATTTGTAATGA 300  
|||||  
Db 390 TGAAGATTTACTCATGAAAGATCCAAACTATACCTAAAGATAGATATTTGTAATGA 449  
QY 301 GGTCAAGGATGATATGTTTATCAAGTATAGTGAAGTATATGATATACCTTAAGTATGTC 360  
|||||  
Db 450 GGTCAAGGATGATATGTTTATCAAGTATAGTGAAGTATATGATATACCTTAAGTATGTC 509  
QY 510 TGCCACGCGGATATGCTGCTCAAAAAGAGAAATCAATGACAAACAGAGCATAG 569  
|||||  
Db 361 TGCCACGCGGATATGCTGCTCAAAAAGAGAAATCAATGACAAACAGAGCATAG 420  
QY 510 TGCCACGCGGATATGCTGCTCAAAAAGAGAAATCAATGACAAACAGAGCATAG 569  
Db 421 TCAACATCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 480  
|||||  
Db 570 TCAACATCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 629  
QY 481 AGAGCGTATATCTACAGATGATGTTTATGCTTAAATGCTCTGATATCATAGAGATAC 540  
|||||  
Db 630 AGAGCGTATATCTACAGATGATGTTTATGCTTAAATGCTCTGATATCATAGAGATAC 689  
QY 541 TGTGATGCTTATATGCTTCTCATGAGATCATTTACATTAATCTCTAAGATGAGTT 600  
|||||  
Db 690 TGTGATGCTTATATGCTTCTCATGAGATCATTTACATTAATCTCTAAGATGAGTT 749  
QY 601 ATCAGCTAGCAGATGTTGCTGTCAGAGAGCTTCTATCTGCTGAGAGAAATCTGTCAA 660  
|||||  
Db 750 ATCAGCTAGCAGATGTTGCTGTCAGAGAGCTTCTATCTGCTGAGAGAAATCTGTCAA 809  
QY 661 TTCAAGAACTATGCGCGACAAATAGAGATTAACATTTCAAGAACTATGCGTTC 720  
|||||  
Db 810 TTCAAGAACTATGCGCGACAAATAGAGATTAACATTTCAAGAACTATGCGTTC 869  
QY 721 TGTAGCAATTCAGAACTACAAATACTAACACAGCAACAGCAACAGCAACAGTCA 780  
|||||  
Db 870 TGTAGCAATTCAGAACTACAAATACTAACACAGCAACAGCAACAGCAACAGTCA 929  
QY 781 AGCAAGTCAAGTATATGATGATGATGCTTCAAGACGCTTCAAGACGCTTTGAG 840  
|||||  
Db 930 AGCAAGTCAAGTATATGATGATGATGCTTCAAGACGCTTCAAGACGCTTTGAG 989  
QY 841 TCAAGCAATGTAGAAATCTAGTGCCTTGTGATTCAGCAACAAATCAAGTCAAC 900

|||||  
Db 990 TCAACGACATGTAGAAATCTGATGGCCCTTGTCTTGATCCAGCAAAATCAAGTGCAG 1049  
QY 901 AGCTAAGAGTGTTCAGCTGGCCACAGAGATCAATTAACACTTCACTCCCTTACTCAAT 960  
Db 1050 AGCTAAGAGTGTTCAGCTGGCCACAGAGATCAATTAACACTTCACTCCCTTACTCAAT 1109  
QY 961 GTCTGAATTTGGAAGAGCAATCGCTGATTTATTTCCCTTCGTTATCGTTCAACCAATG 1020  
Db 1110 GTCTGAATTTGGAAGAGCAATCGCTGATTTATTTCCCTTCGTTATCGTTCAACCAATG 1169  
QY 1021 GGTACACAGATTCAGAGCCACAGAACCAAGTCCACACGAGTCCGAGACCTAGTCCAGG 1080  
Db 1170 GGTACACAGATTCAGAGCCACAGAACCAAGTCCACACGAGTCCGAGACCTAGTCCAGG 1229  
QY 1081 CCGGCAACCTGCACCAATCTTAAATAGACTCAATTCCTCTTGGTTGTGTCAGCTGCT 1140  
Db 1230 CCGGCAACCTGCACCAATCTTAAATAGACTCAATTCCTCTTGGTTGTGTCAGCTGCT 1289  
QY 1141 ACGAAAAGTTGGGGAAGAGATATGTATTCGAAGAAAAGGAGATCTCTGTTATGCTTTGC 1200  
Db 1290 ACGAAAAGTTGGGGAAGAGATATGTATTCGAAGAAAAGGAGATCTCTGTTATGCTTTGC 1349  
QY 1201 GAAAGATTTACATCTGAAAAGCTGTTAAAAATCTTGAAGAGCAAGTTATCAAAACAGAGAG 1260  
Db 1350 GAAAGATTTACATCTGAAAAGCTGTTAAAAATCTTGAAGAGCAAGTTATCAAAACAGAGAG 1409  
QY 1261 TGTTCACACACTTTACTACTCTTAAAAAGAAAATGTCCTCCCTGACCAACAATTTTA 1320  
Db 1410 TGTTCACACACTTTACTACTCTTAAAAAGAAAATGTCCTCCCTGACCAACAATTTTA 1469  
QY 1321 TGATTAAGCATATATATCTGTTAACTGAGGCTCAATAAGCCTTGTGTTGNAATTAAGGCTG 1380  
Db 1470 TGATTAAGCATATATATCTGTTAACTGAGGCTCAATAAGCCTTGTGTTGNAATTAAGGCTG 1529  
QY 1381 TAATTTGATTTCCAGCCTTAGACAATTTATAGAACGCTTGAGATGATGACTAA 1440  
Db 1530 TAATTTGATTTCCAGCCTTAGACAATTTATAGAACGCTTGAGATGATGACTAA 1589  
QY 1441 TAAGAAGAAATTTGATGATTTATTTGGCATTCCTAGCAACCAATTAACCATCCAGAGCG 1500  
Db 1590 TAAGAAGAAATTTGATGATTTATTTATTTGGCATTCCTAGCAACCAATTAACCATCCAGAGCG 1649  
QY 1501 ACTTGGCAAAACCAATTTCTCAATTTAGATTAAGTGAAGAGCAAGTTCTGATTCCTCAAT 1560  
Db 1650 ACTTGGCAAAACCAATTTCTCAATTTAGATTAAGTGAAGAGCAAGTTCTGATTCCTCAAT 1709  
QY 1561 AGCTGATAGTATACACAGTCAAGTGTATCAATTTTGTGTAACATGATTAATCACTGA 1620  
Db 1710 AGCTGATAGTATACACAGTCAAGTGTATCAATTTTGTGTAACATGATTAATCACTGA 1769  
QY 1621 TGAAGAGATGATATGTATAGCGCTCATATGAGGCAATAGTCACTGGATTTGGAAGAGTAG 1680  
Db 1770 TGAAGAGATGATATGTATAGCGCTCATATGAGGCAATAGTCACTGGATTTGGAAGAGTAG 1829  
QY 1681 CCTTTCGTATAGGAAAAAGTTGACAGTCAAGCCTTACTATAAGAAAAAGTATTCCTAAC 1740  
Db 1830 CCTTTCGTATAGGAAAAAGTTGACAGTCAAGCCTTACTATAAGAAAAAGTATTCCTAAC 1889  
QY 1741 TCCATCTCCAGACGAGATGTTAAAGCAATTCACCTGAGATAGTGCACAGCTATTTA 1800  
Db 1890 TCCATCTCCAGACGAGATGTTAAAGCAATTCACCTGAGATAGTGCACAGCTATTTA 1949  
QY 1801 CAATCTGTGAAGAGGGAAGAAAGCAATTCACCTGAGATAGTGCACAGCTATTTA 1860  
Db 1950 CAATCTGTGAAGAGGGAAGAAAGCAATTCACCTGAGATAGTGCACAGCTATTTA 2009  
QY 1861 TACAGTTGAGGTTAAAGAAAGCTATTTGATTTCTCATTAAGAGATATACATATAT 1920  
Db 2010 TACAGTTGAGGTTAAAGAAAGCTATTTGATTTCTCATTAAGAGATATACATATAT 2069  
QY 1921 TAAATTTGCTTGGTTGATGATCACAATCAAGCTCCAAATGGCTATACCTTGGAGA 1980  
|||||

Db 2070 TAAATTTGCTTGGTTGATGATCACAATCAAGAGCTCCAAATGGCTATACCTTGGAGA 2129  
QY 1981 TTTGTTGGCAGATTAAAGTACAGTAGAGAACCCCTGACAGACGTCACATTTCTAATGA 2040  
Db 2130 TTTGTTGGCAGATTAAAGTACAGTAGAGAACCCCTGACAGACGTCACATTTCTAATGA 2189  
QY 2041 TGGATGGGCAATGACAGTAGAGCATGTGTAGGCAAGAAAGACACAGTAGTAATCCAA 2100  
Db 2190 TGGATGGGCAATGACAGTAGAGCATGTGTAGGCAAGAAAGACACAGTAGTAATCCAA 2249  
QY 2101 TAAGAACTTCAAGCCGAGATGAAGACCAAGTAGAGAAACACTGCTGAGCCAGAGTCCC 2160  
Db 2250 TAAGAACTTCAAGCCGAGATGAAGACCAAGTAGAGAAACACTGCTGAGCCAGAGTCCC 2309  
QY 2161 TCAAGTAGACTGAAAAAATAGAGCCCAACCAAGAAAGCAAGAGTTTGGTTGGAA 2220  
Db 2310 TCAAGTAGACTGAAAAAATAGAGCCCAACCAAGAAAGCAAGAGTTTGGTTGGAA 2269  
QY 2221 AGTAACGATTTAGTCTGAAAGCCCAATGCAACAGAAACTCTAGCTGTTACGAATA 2280  
Db 2370 AGTAACGATTTAGTCTGAAAGCCCAATGCAACAGAAACTCTAGCTGTTACGAATA 2429  
QY 2281 TTTGACTCTTCAAAATTAAGTATACCAATAGTATCATGCGCAAGCAAGAAATTAATTG 2340  
Db 2430 TTTGACTCTTCAAAATTAAGTATACCAATAGTATCATGCGCAAGCAAGAAATTAATTG 2489  
QY 2341 GTTGTAAAGAGATATATCTCATCTGTAAGTAAAGAAATTAATAC 2389  
Db 2490 GTTGTAAAGAGATATATCTCATCTGTAAGTAAAGAAATTAATAC 2538  
|||||

RESULT 3  
BD003774 8195 bp DNA linear PAT 31-JAN-2002  
LOCUS  
DEFINITION Polynucleotide of Streptococcus pneumoniae and sequence.  
ACCESSION BD003774.1 GI:18631735  
VERSION JP 2001501833-A/94.  
KEYWORDS unclassified.  
SOURCE unclassified.  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 8195)  
AUTHORS Kunsch, C.A., Choi, G.H., Dillon, P.J., Rosen, C.A., Bara, S.C.,  
Fannon, M. and Dougherty, B.A.  
TITLE Polynucleotide of Streptococcus pneumoniae and sequence  
JOURNAL Patent: JP 2001501833-A 94 13-FEB-2001;  
HUMAN GENOME SCIENCES INC  
COMMENT OS unclassified  
PN JP 2001501833-A/94  
PD 13-FEB-2001  
PE 30-OCT-1997 JP 1998520718  
PR 31-OCT-1996 US 60/029960  
PI CHARLES A KUNSCH, GIL H CHOI, PATRICK J DILLON, CRAIG A ROSEN, PI  
STEVEN C BARASH,  
PI MICHAEL FANNON, BRIAN A DOUGHERTY  
PC C12N15/09, A01K67/027, C07K14/315, C07K16/12, C12N1/15, C12N1/19,  
PC C12N1/21,  
PC C12N5/10, C12P21/02, C12Q1/68, G06F17/30, C12N15/00, C12N5/00, PC  
G06F15/40  
CC Strandedness: Double;  
CC Topology: Linear;  
FH key Location/Qualifiers  
FT source 1. 8195  
Location/Qualifiers  
1. 8195  
/organism="unclassified"  
/db\_xref="taxon:32644"  
BASE COUNT 2688 a 1622 c 1777 g 2105 t 3 others  
ORIGIN  
Query Match 100.0%; Score 2388; DB 6; Length 8195;  
Best local Similarity 100.0%; Pred. No. 0;

Matches 2388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 TTCTTACGATTTGGAGCTGTATCAAGCTTAGACGGTTAAGAAAAATATGTGTTCCCTA 60
    |||||
Db 3053 TTCTTACGATTTGGAGCTGTATCAAGCTTAGACGGTTAAGAAAAATATGTGTTCCCTA 3112

QY 61 TATAGATGAAAAACACGACGCAAAAAACGAGAAATTTGACTCTTGATGAGTTAGCAA 120
    |||||
Db 3113 TATAGATGAAAAACACGACGCAAAAAACGAGAAATTTGACTCTTGATGAGTTAGCAA 3172

QY 121 GCCTGAAGGAATCAATGCTGAGCAAAATCGTCATCAATATACAGACCAAGCTATGTAC 180
    |||||
Db 3173 GCGTGAAGGAATCAATGCTGAGCAAAATCGTCATCAATATACAGACCAAGCTATGTAC 3232

QY 181 TTTCACATGGGACCACTATCTATTATTCAAATGTAAGTTCCTTATGAGCTATCATCAG 240
    |||||
Db 3233 TTTCACATGGGACCACTATCTATTATTCAAATGTAAGTTCCTTATGAGCTATCATCAG 3292

QY 241 TGAAGAAATTAATCATGTAAGATCCAAACTATAGCTAAAAAGATGAGATATTTGTTAATGA 300
    |||||
Db 3293 TGAAGAAATTAATCATGTAAGATCCAAACTATAGCTAAAAAGATGAGATATTTGTTAATGA 3352

QY 301 GGTCAAGGGGGAATGCTTTTCAAGGTAGATGGAATTAATGTTTACTATGTTTAAAGATGC 360
    |||||
Db 3353 GGTCAAGGGGGAATGCTTTTCAAGGTAGATGGAATTAATGTTTACTATGTTTAAAGATGC 3412

QY 361 TGCCCAAGCGGATTAACGTCCGTACAAAGAGAAATCAATCGACAAAAACAGACATAG 420
    |||||
Db 3413 TGCCCAAGCGGATTAACGTCCGTACAAAGAGAAATCAATCGACAAAAACAGACATAG 3472

QY 421 TCACATCTGTGAAGGTGGAACCTCCAAAGAAACGATGTGTGCTGTGCTTGCAAGCTTGCA 480
    |||||
Db 3473 TCACATCTGTGAAGGTGGAACCTCCAAAGAAACGATGTGTGCTGTGCTTGCAAGCTTGCA 3532

QY 481 AGGACGCTACTACTACAGATAGTGTATATCTTAAATGCTTGTGATATCAATAGAGATAC 540
    |||||
Db 3533 AGGACGCTACTACTACAGATAGTGTATATCTTAAATGCTTGTGATATCAATAGAGATAC 3592

QY 541 TGGTGATGTTTATATCGTTCCTCATGAGATCATTTACATTACTTCTTAAGATGAGTT 600
    |||||
Db 3593 TGGTGATGTTTATATCGTTCCTCATGAGATCATTTACATTACTTCTTAAGATGAGTT 3652

QY 601 ATACAGCTAGCGAGTTGGCTGCTCGAGAAAGCCTTCTATCTGTGCTGAGAAATCTGTCAA 660
    |||||
Db 3653 ATACAGCTAGCGAGTTGGCTGCTCGAGAAAGCCTTCTATCTGTGCTGAGAAATCTGTCAA 3712

QY 661 TTCAAGAACTATTCGCCGAAATAATAGCGATTAACACTTCAAGAAACAACTGGGACTTTC 720
    |||||
Db 3713 TTCAAGAACTATTCGCCGAAATAATAGCGATTAACACTTCAAGAAACAACTGGGACTTTC 3772

QY 721 TGTAGCAATCCAGAACTACAAATACTATAACACAGCAAGCAAGCAACTAAGACTCA 780
    |||||
Db 3773 TGTAGCAATCCAGAACTACAAATACTATAACACAGCAAGCAAGCAACTAAGACTCA 3832

QY 781 AGCAAGTCAAAAGTAATGACATTTATAGTCTTGAAGACAGCTCTACAAACTGCTTTGAG 840
    |||||
Db 3833 AGCAAGTCAAAAGTAATGACATTTATAGTCTTGAAGACAGCTCTACAAACTGCTTTGAG 3892

QY 841 TCACAGCAATGTAGATCTATGAGCCCTGTGCTTGTGATCCGACAAATTCACAGTGCAG 900
    |||||
Db 3893 TCACAGCAATGTAGATCTATGAGCCCTGTGCTTGTGATCCGACAAATTCACAGTGCAG 3952

QY 901 AGCTAGAGTGTTCAGAGTGCACACGAGATCATTTACCACTTACCTTACTCTCAAT 960
    |||||
Db 3953 AGCTAGAGTGTTCAGAGTGCACACGAGATCATTTACCACTTACTCTTACTCTCAAT 4012

QY 961 GTCTGAATTTGGAAGAAAGAAATCGCTGTATTTATTCCTTGTATTCGTTATCCAAACAT 1020
    |||||
Db 4013 GTCTGAATTTGGAAGAAAGAAATCGCTGTATTTATTCCTTGTATTCGTTATCCAAACAT 4072

QY 1021 GGTACCAAGATTTCAAGGCGCAAGAACACCAAGTCCACAAACGACATCCGAACTAGTCCAG 1080
    |||||
Db 4073 GGTACCAAGATTTCAAGGCGCAAGAACACCAAGTCCACAAACGACATCCGAACTAGTCCAG 4132
```

```
QY 1081 CCGCAACCTGCACCAATCTTAAATAGACTCAAAATCTCTTGGTGTAGTCACTGCT 1140
    |||||
Db 4133 CCGCAACCTGCACCAATCTTAAATAGACTCAAAATCTCTTGGTGTAGTCACTGCTGCT 4192

QY 1141 ACGAAAAATTTGGGAGGATATGTTTCGAAGAAAAAGGGCATCTCTGTTATGCTTTGTC 1200
    |||||
Db 4193 ACGAAAAATTTGGGAGGATATGTTTCGAAGAAAAAGGGCATCTCTGTTATGCTTTGTC 4252

QY 1201 GAAAGATTTTACCATCTGTAACCTGTTAAATCTTTGAAAGCAAGTATTCAAAACAGAGAG 1260
    |||||
Db 4253 GAAAGATTTTACCATCTGTAACCTGTTAAATCTTTGAAAGCAAGTATTCAAAACAGAGAG 4312

QY 1261 TGTTCACACATTTAATCTGCTAAAAAGAAATGTTGCTCTCTGAGCAAGAAATTTTA 1320
    |||||
Db 4313 TGTTCACACATTTAATCTGCTAAAAAGAAATGTTGCTCTCTGAGCAAGAAATTTTA 4372

QY 1321 TGATTAAGCATATTAATCTGTTAACTGAGGCTCATTAAGCCTTGTGTTGMAATAAGGGTGC 1380
    |||||
Db 4373 TGATTAAGCATATTAATCTGTTAACTGAGGCTCATTAAGCCTTGTGTTGMAATAAGGGTGC 4432

QY 1381 TAATTCGATTTTCCAAAGCCTTAGCAAAATTAATTAAGACCTTGATATGATATGACTAA 1440
    |||||
Db 4433 TAATTCGATTTTCCAAAGCCTTAGCAAAATTAATTAAGACCTTGATATGATATGACTAA 4492

QY 1441 TAAAGAAAAATTTGGTAGATGATTTATTTGCAATTCCTAGACCAATTAACCATCCAGAGCG 1500
    |||||
Db 4493 TAAAGAAAAATTTGGTAGATGATTTATTTGCAATTCCTAGACCAATTAACCATCCAGAGCG 4552

QY 1501 ACTTGGCAAAACCAAAATTCCTCAAAATGAGTATCTGAAGACGAAGTTGATGCTCAATT 1560
    |||||
Db 4553 ACTTGGCAAAACCAAAATTCCTCAAAATGAGTATCTGAAGACGAAGTTGATGCTCAATT 4612

QY 1561 AGCTGATTAAGATATCAACGTCAGATGTTACATTTTGATGAACATGATATTAACAGTGA 1620
    |||||
Db 4613 AGCTGATTAAGATATCAACGTCAGATGTTACATTTTGATGAACATGATATTAACAGTGA 4672

QY 1621 TGAAGGAGATGATATGATTAACGCTCATATGAGGCTAGTCACTGATTTGAAAAAGATAG 1680
    |||||
Db 4673 TGAAGGAGATGATATGATTAACGCTCATATGAGGCTAGTCACTGATTTGAAAAAGATAG 4732

QY 1681 CTTTCTGATTAAGAAAAAGTTGCAAGCTCAAGCTATTAAGAAAAAGATATCTTACC 1740
    |||||
Db 4733 CTTTCTGATTAAGAAAAAGTTGCAAGCTCAAGCTATTAAGAAAAAGATATCTTACC 4792

QY 1741 TCCATCTCCAGAGCGCAATGTTAAAGCAATTCCAACTGGAGATAGTCCAGCAGCTATTTA 1800
    |||||
Db 4793 TCCATCTCCAGAGCGCAATGTTAAAGCAATTCCAACTGGAGATAGTCCAGCAGCTATTTA 4852

QY 1801 CAATCGTGAAGGAGGAAAAAGCAATTCACATGCTGTCGATCCATATATGTTGAGCA 1860
    |||||
Db 4853 CAATCGTGAAGGAGGAAAAAGCAATTCACATGCTGTCGATCCATATATGTTGAGCA 4912

QY 1861 TACAGTTGAGGTTAAAAACGTAATTTGATTAATTCCTCATTAAGATCATTAACATAATAT 1920
    |||||
Db 4913 TACAGTTGAGGTTAAAAACGTAATTTGATTAATTCCTCATTAAGATCATTAACATAATAT 4972

QY 1921 TAAATTTGCTTTGTTTATATATCACACATACAAAGCTTCCAAATGCTATATACCTTGAAGA 1980
    |||||
Db 4973 TAAATTTGCTTTGTTTATATATCACACATACAAAGCTTCCAAATGCTATATACCTTGAAGA 5032

QY 1981 TTTGTTTGGACGATTTAAGTACTAGTAGAAGCAACCCGAGAGAGTCCCATATCTAATGA 2040
    |||||
Db 5033 TTTGTTTGGACGATTTAAGTACTAGTAGAAGCAACCCGAGAGAGTCCCATATCTAATGA 5092

QY 2041 TGGATGGGGAATGCGCAGTAGGATGTTAGGCCAAGAAAGACACAGTAAGATCCAAA 2100
    |||||
Db 5093 TGGATGGGGAATGCGCAGTAGGATGTTAGGCCAAGAAAGACACAGTAAGATCCAAA 5152

QY 2101 TAAGAACTTTCAAAAGCGGATGAAGAGCCAGTAGAGAAACCTGTGAGCCAGAAATGCC 2160
    |||||
Db 5153 TAAGAACTTTCAAAAGCGGATGAAGAGCCAGTAGAGAAACCTGTGAGCCAGAAATGCC 5212
```

OY	2161	TCAAGTAGACGACTGAAAAAGTAGAAGCCCACTCACAAGAAGAGAAAGTTTCCTTGCGGA	2220
Db	5213	TCAAAGTAGAGACTGAAAAGATAGAAAGCCCACTCACAAGAAGAGAAAGTTTCCTTGCGGA	5272
OY	2221	ACTAACGGATTCTAGTCTGTGAAGCCCAATGCACAGAAACTCTAGCTGGTTTACGAATA	2280
Db	5273	ACTAACGGATTCTAGTCTGTGAAGCCCAATGCACAGAAACTCTAGCTGGTTTACGAATA	5332
OY	2281	TTTTGACTCTTCAAAATTATGATTAACAAATATCATGATGCCAGAGAGAAAAATTA	2340
Db	5333	TTTTGACTCTTCAAAATTATGATTAACAAATATCATGATGCCAGAGAGAAAAATTA	5392
OY	2341	GTTGTTAAAGAGAAATACCTTCATCTCTAGTAAAGAAAAAATAAAC	2389
Db	5393	GTTGTTAAAGAGAAATACCTTCATCTCTAGTAAAGAAAAAATAAAC	5441
RESULT 4			
AEO07418/C			
LOCUS	AEO07418	10256 bp	DNA linear BCT 31-AUG-2001
DEFINITION	Streptococcus pneumoniae TIGR4 section 101 of 194 of the complete genome.		
ACCESSION	AEO07418	AEO05672	
VERSION	AEO07418.1	GI:14972649	
KEYWORDS			
SOURCE			
ORGANISM	Streptococcus pneumoniae TIGR4. Streptococcus pneumoniae TIGR4 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.		
REFERENCE	1 (bases 1 to 10256)		
AUTHORS	Tettelin,H., Nelson,K.E., Paulsen,I.T., Eisen,J.A., Read,T.D., Peterson,S., Heidelberg,J., Deboy,R.T., Hatt,D.H., Dodson,R.J., Durkin,A.S., Gwinn,M., Kolonay,J.F., Nelson,W.C., Peterson,J.D., Umayam,L.A., White,O., Salzberg,S.L., Lewis,M.R., Radune,D., Holtzapple,E., Khouri,H., Wolf,A.M., Utterback,T.R., Hansen,C.L., McDonald,L.A., Feldblum,T.V., Anginoli,S., Dickinson,T., Hickey,E.K., Holt,I.E., Loftus,B.J., Yang,F., Smith,H.O., Venter,J.C., Dougherty,B.A., Morrison,D.A., Hollingshead,S.K. and Fraser,C.M.		
TITLE	Complete genome sequence of a virulent isolate of Streptococcus pneumoniae		
JOURNAL	Science 293 (5529), 498-506 (2001)		
MEDLINE	21357209		
PUBMED	11463916		
REFERENCE	2 (bases 1 to 10256)		
AUTHORS	Tettelin,H., Nelson,K.E., Paulsen,I.T., Eisen,J.A., Read,T.D., Peterson,S., Heidelberg,J., Deboy,R.T., Hatt,D.H., Dodson,R.J., Durkin,A.S., Gwinn,M., Kolonay,J.F., Nelson,W.C., Peterson,J.D., Umayam,L.A., White,O., Lewis,M.R., Radune,D., Holtzapple,E., Khouri,H., Wolf,A.M., Utterback,T.R., Hansen,C.L., McDonald,L.A., Feldblum,T.V., Anginoli,S., Geswan,P., Hickey,E.K., Holt,I.E., Loftus,B.J., Ujwal,M.L., Yang,F., Smith,H.O., Venter,J.C., Dougherty,B.A., Morrison,D.A., Hollingshead,S.K. and Fraser,C.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-JUN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
FEATURES	Location/Qualifiers		
SOURCE	1..10256		
	/organism="Streptococcus pneumoniae TIGR4"		
	/strain="TIGR4"		
gene	/db_xref="taxon:170187"		
	complement(100..702)		
	/gene="SP1170"		
CDS	complement(100..702)		
	/gene="SP1170"		
	/note="Identified by Glimmer2; putative"		
	/codon_start=1		
	/transl_table=1		
	/product="hypothetical protein"		
	/protein_id="AAK75278.1"		
	/db_xref="GI:14972650"		
	/translation="MSEVDNEAVNFEFTSDTQLANSIYQSLEKFPDKKNFSGDLIF		
	TWKSPSLVKEBDYIGRDSDQDNRVIGNITFNPTLTNRKYSLNNRNNGCMGDPEHDFE		

DLYIDHVAKYAEQKVVNNIKETYYELPKRALILQENALFREFSNDDPLENNYKLTITWO  
 VSKETPESEMDNPKNKISEKILIFERGSKMINDLSNKKK"  
 cds  
 /gene="SP1171"  
 /complement(760. .1473)  
 /note="identified by match to PFAM protein family HMM PF00702"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hydrolase, haloacid dehalogenase-like family"  
 /protein\_id="AAK75280.1"  
 /db\_xref="GI:14972651"  
 /translation="MEYKFEFLFDLDHTLLDFDAEDVALPQLKEGVADIQAYKDY  
 YPMNKAUKDELKIKYFQELVNTFRSLFHFQGEKDGSLAQROGFYLAOOGTLS  
 GAHDLIDSLIEDSGEYMLYATNGITALIGTGLAOSGLAPYENOVYISOLGTORPDALF  
 YEKIIOGLIAGSEKTEKTLMDISGLADIGGNNAIDITLWNPMLHNHTQAQTYEY  
 SYQDLIDDLCKNILEKTF"  
 cds  
 /complement(1724. .1843)  
 /gene="SP1172"  
 /complement(1724. .1843)  
 /note="identified by GLIMMER2; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="AAK75281.1"  
 /db\_xref="GI:14972652"  
 /translation="MLYAVPEFYFNRSSEITVPLNCSIKTQDCAILALEFYKN"  
 /complement(1933. .2307)  
 /gene="SP1173"  
 /complement(1933. .2307)  
 /note="identified by GLIMMER2; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical protein"  
 /protein\_id="AAK75282.1"  
 /db\_xref="GI:14972653"  
 /translation="MESIGIYVSHKIHAGVVELISKVAKDVPITVYKGTGGGIG  
 NSFEQDVVSENPAIDTLAFLFDLSAKMNLKMTDSDSIINRPVIEGAYNMAA  
 LLQAGAEISLVQIQALAEIKK"  
 cds  
 /complement(2729. .5188)  
 /gene="SP1174"  
 /complement(2729. .5188)  
 /note="similar to GP\_4249624; identified by sequence  
 similarity; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved domain protein"  
 /protein\_id="AAK75283.1"  
 /db\_xref="GI:14972654"  
 /translation="WKINKKYLAGSAVALISCSYELGTRYQAGODKESNRVAYIDG  
 DQAGAPYKLPDEVSCKREGINAEQIVIKITDOGYTSHDHNHYNGVAYIAISE  
 ELKMDPNQKLDSDIWEIKGVYIKVNGYUYLLDAAHANIRKKEIKKOER  
 SNHNSRADMVNAARAAGRRYTDDGYIPNASDIIEPTGDAYIVPGHDHNYIKPML  
 SASLEAAAEAVKKEGILPRSSSSSYNANPQRLSNHNLVTTPYTHQNGSNISL  
 LRELPAKLSRHVESDGLTDPQAQITSRIRAGVAVPHGNHTFIPEDMSLEKRTIA  
 RIIPIRTYSNHWPDVSREEPSQPTEPSPSDQAPAPSNIDELVKEAVKAGDDVY  
 FEENVSAYIPAKDLSETAGIDSKLAKQESLSHKGTRKTDLPSSDRFYNKAYDL  
 LARIHQDLKRGQVDFEALDNLLERLKDVSSDKYKLVEDILAFAPIRHPRLGKP  
 NQIIVYQDELDQVARKGKYTTEDGYIFEDPRDITSDGDAAVYVPMTHSHWIKKDSL  
 EAERAAQAVYKKEGILPRSTDHODSSNTENKAEALYNNYKAAKVPRLDRMYNIQY  
 TVEVYNGSLIIPHYDHNLIKLEWFDGCLTEARGYTLEDLATVAKTYVEHPERHS  
 DNGEFAADHVAQRNKGADINQEQEPSEKRPQTEKEEETPREEKROSPESSPPT  
 DEEPSEPESEEPQVETEKEVKEAREADELDLGIQDPIISNNAKETLTGLKNMLFGT  
 QDNNTIMAEAKLALLKESK"  
 cds  
 /complement(5346. .7754)  
 /gene="SP1175"  
 /complement(5346. .7754)  
 /note="similar to GP\_4249624; identified by sequence

```
similarity: putative"
/codon_start=11
/transl_table=11
/product="conserved domain protein"
/protein_id="AAK75284.1"
/db_xref="GI:14972655"
/translaction="MILSVCSYELGLYQARTVKNRVSIDKQATORTENLTJPEV
SKREDINADIVIKITIDOGVYSHGDHYHNGKVPYD1ISELLMKPKNYLKED
IYNEVGVYIKDKKYIYLKDAHADVNRTEINRPOKQESORREGTPENDAV
AARSGRYTTDGGYIFENASDIIEDTGDVIVPHGHYHYPENELISAGLAEKFL
SRGNTLSNRTRYRONSDMSTRNWPYSNPECTTNTNNSNTSOSQSDIDSL
LKOLYLPISQRHVESDGLVFPDPAQITTSRARGVAPHGDHFIYISQSELEERIA
RIIPKRYRSNHWPDSRPEQSPQPTPEPSPPQAPNPKIDISNSLSQLVKVEBG
YVEEEKISRYVFAKDLPSYTKNLESKLSKQESVSHLTAKENYAPRDOFVYDKAY
NLTLEAKALFEKNGRNSDFOALDKLERLENDSTKEKLVDDLAFAPITPERLG
KPSQLEYEDEVRIALADKYTSDGYIFEDEDIISDGDVAVPHGHSHHIGDS
LSDKKEVAOAYTEKEGTLIPSPADYKANPCDSAAIYNRYKGEKRIPIYLPVW
BHYEVKNKNLIIIPKDHYNHNIKFAFDDHTYKAPNGYILEDLFIATIKTYVEHPDERP
HSDMGVNASSEHYLGKRDHSEDPKNFKADEEVEETPEPEVPOYTEKREVAQLKEA
EVLAKVYDSSLKANATELTLAGLRNNLTQIMDNNSIMAEKLLALLKGNPSYSVK
EKIN"
complement(8049..9782)
/gene="SP1176"
complement(8049..9782)
/gene="SP1176"
/transl_table=11
/codon_start=11
/product="phosphoenolpyruvate-protein phosphotransferase"
/protein_id="AAK75285.1"
/db_xref="GI:14972656"
/translaction="MTEMLKGIASDGVAVAKAYLLVOPDSFETITVEDTNAEARL
DIALQASODELSVIREKAVGLTGEBAQVDAHLVADPEMSIQIKETIRAKVAAE
AGLEKTYDMETITEGEMEDNPYMOERADIRDTYKLVANLKKIPNPSISEEIVY
IAHDLIPSDTAOLDKNFKAPVNTNIGSRSHSIAKRTLEIAVGTNNITETLVKGD
IIAVNGITGEVINIPTDEQAEKAGEARAKAKAMALLKDAQITYTAQKHELAN
IGTPKVEYVNNNGEAVGLRTEFLYMSQDEPTDEQYEAKVALEENNGPVVVR
TMDIGDKELPYFDMPEKNPFLGFRLRISSETGDAMFRIOIRALLASVNGQIRI
MEPVVALKEFRKAAVPDEBEKANLAEQAVANDIOVIMIEIPAAAMLADFAKEV
DEFSISTNDLQITMAADRMNEQVSYLYQPYNSIIRLNNVKAHAAGKMGKGE
MAGDOQAVPLVGMGIDFESMSATSVLRSLMKKIDTAKMEYANRALTECSTMEV
LELOKRYVNF"
complement(9788..10051)
/gene="SP1177"
complement(9788..10051)
/gene="SP1177"
/codon_start=11
/transl_table=11
/product="phosphocarrier protein Hpr-"
/protein_id="AAK75286.1"
/db_xref="GI:14972657"
/translaction="MASKDFHVVAFTGIAHARPATLIVOTASKFASDITLEYKGSVNL
KSIIMGVSGVGGADVTISABGADADDAIIMISEMKEBGLA"
BASE COUNT      2842 a      2089 c      1850 g      3475 t
ORIGIN
Query Match      100.0%; Score 2388; DB 1; Length 10256;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTCTTAGAATTGGGACTGTATCAAGCTAGAACGGTTAAGGAAATATATGCTGTTTCTTA 60
|||||
Db 7737 TTCTTAGAATTGGGACTGTATCAAGCTAGAACGGTTAAGGAAATATATGCTGTTTCTTA 7678
OY 61 TATAGATGAAAAACGACGACGAAAAAAGAGAAATTTGACCTCCTATGAGGTTACCA 120
|||||
Db 7677 TATAGATGAAAAACGACGACGAAAAAAGAGAAATTTGACCTCCTATGAGGTTACCA 7618
OY 121 GCGTGAAGAAATCAATGCTGACCAATGCTATCAAGATACAGACCAAGGCTATGTCAC 180
|||||
Db 7617 GCGTGAAGAAATCAATGCTGACCAATGCTATCAAGATACAGACCAAGGCTATGTCAC 7558
OY 181 TTTCACATGGGACACATATATTTACATGTTAGGTTCTTATGACCTATCATCAG 240
|||||
```

```
Db 7557 TTTCACATGGGACACATATATTTACAAATGTTAGGTTCTTATGACCTATCATCAG 7498
OY 241 TGAAGATTTACTCATGAAAGATCCAACTATPAGCTAAAGATGAGGATATTTGTAATGA 300
|||||
Db 7497 TGAAGATTTACTCATGAAAGATCCAACTATPAGCTAAAGATGAGGATATTTGTAATGA 7438
OY 301 GGTCAAGGTTGAGATATGTTATCAAGTAGATGGAATAATCTATGTTTACCTTAAGATGC 360
|||||
Db 7437 GGTCAAGGTTGAGATATGTTATCAAGTAGATGGAATAATCTATGTTTACCTTAAGATGC 7378
OY 361 TGCCCAAGCGGATATACGTCCTGACAAAAGAGAAATCAATCGACAAAACAGAGCATAG 420
|||||
Db 7377 TGCCCAAGCGGATATACGTCCTGACAAAAGAGAAATCAATCGACAAAACAGAGCATAG 7318
OY 421 TCAACATCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 480
|||||
Db 7317 TCAACATCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7258
OY 481 AGACGCTATACATACAGATGATGTTATATCTTTAATGCTTGTATATCATAGAGATAC 540
|||||
Db 7257 AGACGCTATACATACAGATGATGTTATATCTTTAATGCTTGTATATCATAGAGATAC 7198
OY 541 TGGTGATGCTTATATGCTTCTCATGAGATCATTTACCATTTACTTCTTAAGATGATT 600
|||||
Db 7197 TGGTGATGCTTATATGCTTCTCATGAGATCATTTACCATTTACTTCTTAAGATGATT 7138
OY 601 ATCAGCTAGGAGATGCTGCTGTCGACAGAGCTTCCTATGCTGTCGACAGAAATGTCGAAA 660
|||||
Db 7137 ATCAGCTAGGAGATGCTGCTGTCGACAGAGCTTCCTATGCTGTCGACAGAAATGTCGAAA 7078
OY 661 TTCAGAACCTATTCGCCGACAAAATATGCAATACCTTCAAGAAACAACCTGGATACCTTC 720
|||||
Db 7077 TTCAGAACCTATTCGCCGACAAAATATGCAATACCTTCAAGAAACAACCTGGATACCTTC 7018
OY 721 TGTATAGCAATCCAGAACTCAAAATACTAACAAGCAACAGCAACAGCACTAACAGTCA 780
|||||
Db 7017 TGTATAGCAATCCAGAACTCAAAATACTAACAAGCAACAGCAACAGCACTAACAGTCA 6958
OY 781 AGCAGTCAAGTATATGCAATGATGCTGTTGAAACACGCTGACAACTGCTTGGAG 840
|||||
Db 6957 AGCAGTCAAGTATATGCAATGATGCTGTTGAAACACGCTGACAACTGCTTGGAG 6898
OY 841 TCAACGACATGTAGAAATCTGATGCGCTTGTCTTGTATCCAGCAAAATCAAGTCGAC 900
|||||
Db 6897 TCAACGACATGTAGAAATCTGATGCGCTTGTCTTGTATCCAGCAAAATCAAGTCGAC 6838
OY 901 AGCTAGAGGTGTGACGTGCCACAGAGATCATTTACCACTTCATCCCTACTCTCAAT 960
|||||
Db 6837 AGCTAGAGGTGTGACGTGCCACAGAGATCATTTACCACTTCATCCCTACTCTCAAT 6778
OY 961 GTCGATTTGGAAGAACGATGCTGATTTATTCCTCCCTGTTATGCTCAAACTG 1020
|||||
Db 6777 GTCGATTTGGAAGAACGATGCTGATTTATTCCTCCCTGTTATGCTCAAACTG 6718
OY 1021 GGTACGATTTCAAGGCGCAGAACACCAAGTCCACACGACTCCGAACTAGTCCAGG 1080
|||||
Db 6717 GGTACGATTTCAAGGCGCAGAACACCAAGTCCACACGACTCCGAACTAGTCCAGG 6658
OY 1081 CCGGCAACCTGACCAAAATCTTAAATATGACTCAAAATCTTCTTGTGTTGTAAGCTGCT 1140
|||||
Db 6657 CCGGCAACCTGACCAAAATCTTAAATATGACTCAAAATCTTCTTGTGTTGTAAGCTGCT 6598
OY 1141 ACGAAAGTTGGGGAAGATATGATTCGAAGAAAGGGGATCTCTGCTATGCTTGGC 1200
|||||
Db 6597 ACGAAAGTTGGGGAAGATATGATTCGAAGAAAGGGGATCTCTGCTATGCTTGGC 6538
OY 1201 GAAAGATTTACATCTGAAAATCTTAAATACTTGAAGCAAGTTATCAAAACAGAGAG 1260
|||||
Db 6537 GAAAGATTTACATCTGAAAATCTTAAATACTTGAAGCAAGTTATCAAAACAGAGAG 6478
OY 1261 TGTTTACACACTTTTAACCTGCTAATAAAGAAATGTTGCTCTGTCGACCAAGAAATTTTA 1320
|||||
Db 6477 TGTTTACACACTTTTAACCTGCTAATAAAGAAATGTTGCTCTGTCGACCAAGAAATTTTA 6418
```



QY	1321	TGATAAAGCATATATCTGTAACTAGAGGCTCAATAAGCCTGTGTGNAATTAAGGGTCG	1380
Db	6417	TGATTAAGCATATATCTGTAACTAGAGGCTCAATAAGCCTGTGTGNAATTAAGGGTCG	6358
QY	1381	TAATTCGATTTCCAAACCTTTAGACAATAATTATAGAACGCTTGATGATGAAAGCTACTA	1440
Db	6357	TAATTCGATTTCCAAACCTTTAGACAATAATTATAGAACGCTTGATGATGAAAGCTACTA	6298
QY	1441	TAAAGAAAATTGGTAGATCATTTATTTGGCANTTCTAGCACAATTTACCATCCAGACG	1500
Db	6297	TAAAGAAAATTGGTAGATCATTTATTTGGCANTTCTAGCAGACCAATTTACCATCCAGACG	6238
QY	1501	ACTTGGCAAAACCAATTTCTGAAATTTAGATACTGTAAGACGAAGTTCGATTTGGCTCAAT	1560
Db	6237	ACTTGGCAAAACCAATTTCTGAAATTTAGATACTGTAAGACGAAGTTCGATTTGGCTCAAT	6178
QY	1561	AGCTGATAGATATACAGCTCAGATGGTTACATTTTGTATGAACATGATATATACGTGA	1620
Db	6117	TGAAGAGATGCAATATGTAAGCCTCATATATGGGCCATACCTCATGATTTGGAAAAAGATG	1680
QY	1681	CCTTTCGATTAAGGAAAAAGTTGCAGCTCAAGCCTATACTAAGAAGAAAAGGTATCCTACC	1740
Db	6057	CCTTTCGATTAAGGAAAAAGTTGCAGCTCAAGCCTATACTAAGAAGAAAAGGTATCCTACC	5998
QY	1741	TCCATCTCCACAGCAGCATGTTTAAACCAATTCACACTGAGATATGTCAGACAGCTATTTA	1800
Db	5997	TCCATCTCCACAGCAGCATGTTTAAACCAATTCACACTGAGATATGTCAGACAGCTATTTA	5938
QY	1801	CAATCGTGTGAAGGGGAAAAAGCAATTCACCTCGTTCACACTCCATATATGTTGGTAGCA	1860
Db	5937	CAATCGTGTGAAGGGGAAAAAGCAATTCACCTCGTTCACACTCCATATATGTTGGTAGCA	5878
QY	1861	TACAGTTGAGGTTTAAAAACGGTAAATTTGATATTTCTCTCAATAAGATCATACCATATAT	1920
Db	5877	TACAGTTGAGGTTTAAAAACGGTAAATTTGATATTTCTCTCAATAAGATCATATATAT	5818
QY	1921	TAAATTTGCTGGTTGTTTATGATATACACATACAAAGCTCCAAATGGCTTATACCTTGAAGA	1980
Db	5817	TAAATTTGCTGGTTGTTTATGATATACACATACAAAGCTCCAAATGGCTTATACCTTGAAGA	5758
QY	1981	TTTGTTCGACGATTAAGTACTAGTACGTAGAACAACCTCTGACGAGCTCCACATTTCTATGA	2040
Db	5757	TTTGTTCGACGATTAAGTACTAGTACGTAGAACAACCTCTGACGAGCTCCACATTTCTATGA	5698
QY	2041	TGGATGGGGCAATCCCGTAGCATGTGTTTAGGCAAGAAAGCCACACTGSAAGATCCAAA	2100
Db	5697	TGGATGGGGCAATCCCGTAGCATGTGTTTAGGCAAGAAAGCCACACTGSAAGATCCAAA	5638
QY	2101	TAAAGACTTCAAAAGCGGATGAAGAGCCAGTAGAGAAACACCTGCTGAAGCCAGAAAGTCCC	2160
Db	5637	TAAAGACTTCAAAAGCGGATGAAGAGCCAGTAGAGAAACACCTGCTGAAGCCAGAAAGTCCC	5578
QY	2161	TCAAGTGAAGACATGAAAAAGTAAACCCCACTCAAAAGAGAGAAAGTTTTCCTTGGGAA	2220
Db	5577	TCAAGTGAAGACATGAAAAAGTAAACCCCACTCAAAAGAGAGAAAGTTTTCCTTGGGAA	5518
QY	2221	AGTAAAGGATTTAGTCTGTAAGACCAATCAACAGACAAACTAGCTGAGTTCACGAAATTA	2280
Db	5517	AGTAAAGGATTTAGTCTGTAAGACCAATCAACAGACAAACTAGCTGAGTTCACGAAATTA	5458
QY	2281	TTTGACTCTTCAAAATTAATGATTAACAATAGTATCATGTGAGAGACGAGAAAAATTAAGTTC	2340
Db	5457	TTTGACTCTTCAAAATTAATGATTAACAATAGTATCATGTGAGAGACGAGAAAAATTAAGTTC	5398
QY	2341	GTTCCTTAAAGAGAGTAATCTCTTCACTCTGATAGTAAAGAAAAAATAAAC	2389
Db	5397	GTTCCTTAAAGAGAGTAATCTCTTCACTCTGATAGTAAAGAAAAAATAAAC	5349

[illegible]





QY 301 GGTCAAGGTGATATGTTATCAAGTAGATGGAATAACTATGTTTACCTTAAGATGC 360  
|||||  
Db 7508 GGTCAAGGTGATATGTTATCAAGTAGATGGAATAACTATGTTTACCTTAAGATGC 7449  
QY 361 TGGCCACGGGGAATACGTCCTCAAAAGGGAATCAATCGACAAAACAGAGATAG 420  
|||||  
Db 7448 TGGCCACGGGGAATACGTCCTCAAAAGGGAATCAATCGACAAAACAGAGATAG 7389  
QY 421 TCACATCTGTAAGGTGGAACCTCAAGAAACGATGGTGGTGGCTTGGACGTTGCGA 480  
|||||  
Db 7388 TCACATCTGTAAGGTGGAACCTCAAGAAACGATGGTGGTGGCTTGGACGTTGCGA 7329  
QY 481 AGGACGCTACTACAGATGATGTTATATCTTTAAATGCTTGTATATCATAGAGATAC 540  
|||||  
Db 7328 AGGACGCTACTACAGATGATGTTATATCTTTAAATGCTTGTATATCATAGAGATAC 7269  
QY 541 TGGTATGCTTATATGCTTCCATGAGATCATTTACCATTAATTCCTTAAGAAATAGTT 600  
|||||  
Db 7268 TGGTATGCTTATATGCTTCCATGAGATCATTTACCATTAATTCCTTAAGAAATAGTT 7209  
QY 601 ATCAGCTAGCAGATTGGCTGCTCAGAGCCCTTCTATCTGTCGAGAAATCTGTCAA 660  
|||||  
Db 7208 ATCAGCTAGCAGATTGGCTGCTCAGAGCCCTTCTATCTGTCGAGAAATCTGTCAA 7149  
QY 661 TTCAGAACCTTATCGCCGACAAAATAGCGATTAACACTTCAAGAACAACTGGGTACTTC 720  
|||||  
Db 7148 TTCAGAACCTTATCGCCGACAAAATAGCGATTAACACTTCAAGAACAACTGGGTACTTC 7089  
QY 721 TGTAAACCATCCAGAAACATAAATACTAACACAGCAACAGAAACACTAAGCTGA 780  
|||||  
Db 7088 TGTAAACCATCCAGAAACATAAATACTAACACAGCAACAGAAACACTAAGCTGA 7029  
QY 781 AGCAACTCAAGTAATGATGATGATGCTCTTGAAGACGCTCTACAAATGCTTTGAG 840  
|||||  
Db 7028 AGCAACTCAAGTAATGATGATGATGCTCTTGAAGACGCTCTACAAATGCTTTGAG 6969  
QY 841 TCACAGACATGTAAGATCTGATGGCTTGTCTTGTATCCAGACAAATACAAAGTGAAC 900  
|||||  
Db 6968 TCACAGACATGTAAGATCTGATGGCTTGTCTTGTATCCAGACCAATATCAAAAGTGAAC 6909  
QY 901 AGCTAGAGGTGGACGTCACAGGAGATGATGCTTACCACTTACCTTACTCTCAAT 960  
|||||  
Db 6908 AGCTAGAGGTGGACGTCACAGGAGATGATGCTTACCACTTACCTTACTCTCAAT 6849  
QY 961 GTCGTAAATGGAAGAACGATGCTGATTTATCCCTTTCGTTATCGTTCAAACTTGG 1020  
|||||  
Db 6848 GTCGTAAATGGAAGAACGATGCTGATTTATCCCTTTCGTTATCGTTCAAACTTGG 6789  
QY 1021 GGTACCAATTCAGAGCCAGAACACCAAGTCACAAACGATCGGAACTAGTCCAG 1080  
|||||  
Db 6788 GGTACCAATTCAGAGCCAGAACACCAAGTCACAAACGATCGGAACTAGTCCAG 6729  
QY 1081 CCGGCAACCTGGACCAAAATCTTAAATATAGATCAATTTCTCTTGGTATAGTACGTG 1140  
|||||  
Db 6728 CCGGCAACCTGGACCAAAATCTTAAATATAGATCAATTTCTCTTGGTATAGTACGTG 6669  
QY 1141 ACGAAAAAGTTGGGAGATATGATTCGAGAAAAAGGCATCTCGTTATGCTTTTGC 1200  
|||||  
Db 6668 ACGAAAAAGTTGGGAGATATGATTCGAGAAAAAGGCATCTCGTTATGCTTTTGC 6609  
QY 1201 GAAAGATTTACCATCTGAAACCTGTAAATAATCTTGAAGACAGTTTCAAAACAGAG 1260  
|||||  
Db 6608 GAAAGATTTACCATCTGAAACCTGTAAATAATCTTGAAGACAGTTTCAAAACAGAG 6549  
QY 1261 TGTTCACACACTTAACTGCTTAAAGAAATGTTGCTCCGAGACCAAGATTTTA 1320  
|||||  
Db 6548 TGTTCACACACTTAACTGCTTAAAGAAATGTTGCTCCGAGACCAAGATTTTA 6489  
QY 1321 TGAATAAGCATATATCTGTTAACTAGAGCTCATAAAGCCTGTTGNAATTAAGGCTG 1380  
|||||  
Db 6488 TGAATAAGCATATATCTGTTAACTAGAGCTCATAAAGCCTGTTGNAATTAAGGCTG 6429  
QY 1381 TAAATCTGATTTCCAAAGCCTTAGACAAATTTATTAGAAGCCTGATGATGAATGCACTAA 1440

Db 6428 TAAATCTGATTTCCAAAGCCTTAGACAAATTTATTAGAACCTTGAAGATGAATGCACTAA 6369  
QY 1441 TAAAGAAAAATGGTATGATGATTTATTTGGCATTTCCAGACCAATTAACCATCCAGAGC 1500  
|||||  
Db 6368 TAAAGAAAAATGGTATGATGATTTATTTGGCATTTCCAGACCAATTAACCATCCAGAGC 6309  
QY 1501 ACTTGGCAACCAAAATTTCCAAATTTAGTATGATGTAAGAGAGAAATGCTGATTTGCTCAAT 1560  
|||||  
Db 6308 ACTTGGCAACCAAAATTTCCAAATTTAGTATGATGTAAGAGAGAAATGCTGATTTGCTCAAT 6249  
QY 1561 AGCTGATTAAGTATACAGCTCAAGATGTTACATTTTGTATGAACATGATTAATACGTGA 1620  
|||||  
Db 6248 AGCTGATTAAGTATACAGCTCAAGATGTTACATTTTGTATGAACATGATTAATACGTGA 6189  
QY 1621 TGAAGAGATGATATGATTAAGCCTCATATGGGCCATCTACTGATGTAAGAAAGATG 1680  
|||||  
Db 6188 TGAAGAGATGATATGATTAAGCCTCATATGGGCCATCTACTGATGTAAGAAAGATG 6129  
QY 1681 CCTTCTGATTAAGAAAAAGTTGACAGCTCAAGCCTTATCTAAAGAAAAAGTATCTTACC 1740  
|||||  
Db 6128 CCTTCTGATTAAGAAAAAGTTGACAGCTCAAGCCTTATCTAAAGAAAAAGTATCTTACC 6069  
QY 1741 TCATCTCCAGACGATGTTTAAAGCAAAATCCAACTGAGATAGTGCAGACTATTTA 1800  
|||||  
Db 6068 TCATCTCCAGACGATGTTTAAAGCAAAATCCAACTGAGATAGTGCAGACTATTTA 6009  
QY 1801 CAATGCTGTAAGAAAAAGGAAAAAGAAATTCACCTGCTGACATCCATATATGTTGAGCA 1860  
|||||  
Db 6008 CAATGCTGTAAGAAAAAGGAAAAAGAAATTCACCTGCTGACATCCATATATGTTGAGCA 5949  
QY 1861 TACAGTTGAGGTTTAAAGAAAGTGAATTTGATTTATCTCTTAAGATCATTAACATTAAT 1920  
|||||  
Db 5948 TACAGTTGAGGTTTAAAGAAAGTGAATTTGATTTATCTCTTAAGATCATTAACATTAAT 5889  
QY 1921 TAAATTTGCTGTTGATGATGATCAGCATTAAGAGTCCAAATGCTATACCTTGGAGA 1980  
|||||  
Db 5888 TAAATTTGCTGTTGATGATGATCAGCATTAAGAGTCCAAATGCTATACCTTGGAGA 5829  
QY 1981 TTTGTTGGCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
|||||  
Db 5828 TTTGTTGGCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5769  
QY 2041 TGGATGAGGCAATGCCAGTGAAGCATGTTTGAAGCAAGAAACCAAGTGAAGATCCAA 2100  
|||||  
Db 5768 TGGATGAGGCAATGCCAGTGAAGCATGTTTGAAGCAAGAAACCAAGTGAAGATCCAA 5709  
QY 2101 TAAAGATTTCAAGCGGATGGAAGAGCCAGTGAAGAAACACTGCTGAGCCAGAACTCC 2160  
|||||  
Db 5708 TAAAGATTTCAAGCGGATGGAAGAGCCAGTGAAGAAACACTGCTGAGCCAGAACTCC 5649  
QY 2161 TCAAGTAGAGACTGAAAAAGTGAAGAGCCCACTCAAGAAAGAGAAATTTGCTGCGAA 2220  
|||||  
Db 5648 TCAAGTAGAGACTGAAAAAGTGAAGAGCCCACTCAAGAAAGAGAAATTTGCTGCGAA 5589  
QY 2221 AGTAAGGATTTAGTCTGAAAAAGCAATGCAACAGAACTAGCTGTTTGAAGAAATTA 2280  
|||||  
Db 5588 AGTAAGGATTTAGTCTGAAAAAGCAATGCAACAGAACTAGCTGTTTGAAGAAATTA 5529  
QY 2281 TTTGACTCTTCAAAATTTATGATTAACATAGTATCATGAGAGAGCAAGAAATTTACTGC 2340  
|||||  
Db 5528 TTTGACTCTTCAAAATTTATGATTAACATAGTATCATGAGAGAGCAAGAAATTTACTGC 5469  
QY 2341 GTTGTAAAGAGAGATATCTTCAATCTGATGATGATGATGATGATGATGATGATGATG 2389  
|||||  
Db 5468 GTTGTAAAGAGAGATATCTTCAATCTGATGATGATGATGATGATGATGATGATGATG 5420

RESULT 6  
SPNEU1915 20035 bp DNA linear HTG 11-Jul-2001  
LOCUS SPNEU1915  
DEFINITION Streptococcus pneumoniae clone G54, \*\*\* SEQUENCING IN PROGRESS \*\*\*.  
In Ordered pieces.

ACCESSION AL449937  
VERSION AL449937.1 GI:11545162  
KEYWORDS HTG; HTGS\_PHASE2.  
SOURCE Streptococcus pneumoniae.  
ORGANISM Streptococcus pneumoniae  
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
Streptococcus.  
REFERENCE 1 (bases 1 to 20035)  
AUTHORS Dopazo, J., Mendoza, A., Herrero, J., Caldera, F., Humbert, Y.,  
Friedli, L., Guerrier, M., Grand-Schenk, E., Gandin, C., de  
Francesco, M., Polissi, A., Buell, G., Feger, G., Garcia, E., Peitsch, M.  
and Garcia-Bustos, J.F.  
Annotated draft genomic sequence from a Streptococcus pneumoniae  
type 19F clinical isolate  
JOURNAL Microb. Drug Resist. 7 (2), 99-125 (2001)  
MEDLINE 21335329  
PUBMED 11442348  
2 (bases 1 to 20035)  
Dopazo, J., Mendoza, A., Herrero, J., Caldera, F., Polissi, A.,  
Humbert, Y., Friedli, L., Guerrier, M., Grand-Schenk, E., Gandin, C., de  
Francesco, M., Buell, G., Feger, G., Garcia, E., Peitsch, M. and  
Garcia-Bustos, J.F.  
Direct Submission  
JOURNAL Submitted (31-OCT-2000) Research Department, Glaxo Wellcome, S.A.,  
Severo Ochoa 2, 28760 Tres Cantos, SPAIN  
COMMENT \* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
FEATURES  
source 1. 20035  
/organism="Streptococcus pneumoniae"  
/serotype="19F"  
/db\_xref="taxon:1313"  
/clone="G54"  
BASE COUNT 6369 a 3635 c 4108 g 5909 t 14 others  
ORIGIN  
Query Match 97.5%; Score 2328.2; DB 2; Length 20035;  
Best Local Similarity 98.5%; Pred. No. 0;  
Matches 2353; Conservative 6; Mismatches 30; Indels 1; Gaps 1;  
QY 1 TTCTTACAGATTGGAGCTGTATCAAGCTAGAACGGTTAAGGAAATATATCGTTTCTTA 60  
DB 17070 TTCTTACAGATTGGAGCTGTATCAAGCTAGAACGGTTAAGGAAATATATCGTTTCTTA 17129  
QY 61 TATATGATGGAACACACGACGCGCAAAACGAGAGATTGACTCTCTATGAGGTTACCAA 120  
DB 17130 TATATGATGGAACACACGACGCGCAAAACGAGAGATTGACTCTCTATGAGGTTACCAA 17189  
QY 121 GCGTGAAGATCAATCAATCTGAGCAATCGTCATCAAGATTAACGACGAGCTATGTCAC 180  
DB 17190 GCGTGAAGATCAATCTGAGCAATCGTCATCAAGATTAACGACGAGCTATGTCAC 17249  
QY 181 TTACATGGGACACCATATCATTTATCAATGTTAGGTTCTTATGACGCTATCATCAG 240  
DB 17250 TTACATGGGACACCATATCATTTATCAATGTTAGGTTCTTATGACGCTATCATCAG 17309  
QY 241 TGAAGATTTACTATGAAAGATCAAACTATAGCTTAAAGATAGAGATATGTTAATGA 300  
DB 17310 TGAAGATTTACTATGAAAGATCAAACTATAGCTTAAAGATAGAGATATGTTAATGA 17369  
QY 301 GGTCAGAGGTGATATATGTATCAAGTATGAGTGAATATGATATGTTTACCTTAAGGATGC 360  
DB 17370 GGTCAGAGGTGATATATGTATCAAGTATGAGTGAATATGATATGTTTACCTTAAGGATGC 17429  
QY 361 TGGCCACGCGGATATAGCTCGCTACAAAGAGAAATCAATCGACAAAAAAGACAGCATAG 420  
DB 17430 TGGCCACGCGGATATAGCTCGCTACAAAGAGAAATCAATCGACAAAAAAGACAGCATAG 17489  
QY 421 TCAACATCGTGAAGGTGGAAGTCCAAAGAAAGATGCTGCTGCTGGCAGCTTGCGCA 480  
DB 17490 TCAACATCGTGAAGGTGGAAGTCCAAAGAAAGATGCTGCTGCTGGCAGCTTGCGCA 17549

QY 481 AGACAGCTATACATACAGATGATGTTATATCTTAATGCTTGTGATATCATAGAGATAC 540  
DB 17550 GGGAGCGTACACACACACATATATGTTATATCTTTAATGCTTGTGATATCATAGAGAC 17609  
QY 541 TGTGATGCTTATATATGCTTCTTCTGAGATCATTTACATTTACATTTCTAAGATGAGTT 600  
DB 17610 GGTGATGCTTATATATGCTTCTTCTGAGATCATTTACATTTACATTTCTAAGATGAGTT 17669  
QY 601 ATCAGCTAGCGAGTTGGCTGCTGAGAGGCTTCTTATCTGTTGAGAGAAATCTGTCAAA 660  
DB 17670 ATCAGCTAGCGAGTTGGCTGCTGAGAGGCTTCTTATCTGTTGAGAGAAATCTGTCAAA 17729  
QY 661 TTCAAGACCTATGCGGACAAATATAGCATTAACCTTCAAGAACAAATCGGATCTTC 720  
DB 17730 TTCAAGACCTATGCGGACAAATATAGCATTAACCTTCAAGAACAAATCGGATCTTC 17789  
QY 721 TGTATGCAATCCAGGAATCAATCAATATCAACACAGCAACAACTAAGACTCA 780  
DB 17790 TGTATGCAATCCAGGAATCAATCAATATCAACACAGCAACAACTAAGACTCA 17849  
QY 781 AGCAAGTCAAAAGTATATGATATGATAGTCTCTTGAACAGCTGTACAACTGCTTTGAG 840  
DB 17850 ARCAAGTCAAAAGTATATGATATGATAGTCTCTTGAACAGCTGTACAACTGCTTTGAG 17909  
QY 841 TCAACGATGTATGATATCTGATGCTGCTTGTGATCCAGCAACAAATCCAACTGCAAC 900  
DB 17910 TCAACGATGTATGATATCTGATGCTGCTTGTGATCCAGCAACAAATCCAACTGCAAC 17969  
QY 901 AGCTAGAGGTGCTTGCAGTGCACACGAGATCATTTACACTTCACTTCTTACTCTCAAT 960  
DB 17970 AGCTAGAGGTGCTTGCAGTGCACACGAGATCATTTACACTTCACTTCTTACTCTCAAT 18029  
QY 961 GTCTGAATTTGGAAGAACGAATCGCTGATATATTCCTTCTGTTATGTTCAACCAATG 1020  
DB 18030 GTCTGAATTTGGAAGAACGAATCGCTGATATATTCCTTCTGTTATGTTCAACCAATG 18089  
QY 1021 GGTACAGATTCAAGGGCGACAGAACCAAGTCCCAACCGGACCTGAGCTAGTCCAGG 1080  
DB 18090 GGTACAGATTCAAGGGCGACAGAACCAAGTCCCAACCGGACCTGAGCTAGTCCAGG 18149  
QY 1081 CCGCAACCTGACCAATATCTTAAATATGACTCAATTTCTTCTTGTGATGAGTGTGCT 1140  
DB 18150 CCGCAACCTGACCAATATCTTAAATATGACTCAATTTCTTCTTGTGATGAGTGTGCT 18209  
QY 1141 ACGAAAGTTGGGGAAGATATGTTTCGAAGAAAGGCAATCTCTGTTATGCTTTGCT 1200  
DB 18210 ACGAAAGTTGGGGAAGATATGTTTCGAAGAAAGGCAATCTCTGTTATGCTTTGCT 18269  
QY 1201 GAAAGATTTACCATCTGAACCTGTTAAATCTTGAAGAACAGTATCAAAACAAAGAGAG 1260  
DB 18270 GAAAGATTTACCATCTGAACCTGTTAAATCTTGAAGAACAGTATCAAAACAAAGAGAG 18329  
QY 1261 TGTTCACACACTTAACTGCTTAAAGAAATGTTGCTCTCTGAGCAAGAAATTTTA 1320  
DB 18330 TGTTCACACACTTAACTGCTTAAAGAAATGTTGCTCTCTGAGCAAGAAATTTTA 18389  
QY 1321 TGATTAAGCATATATCTGTTTAACTGAGGCTCATTAAGCCTTGTGTTTAAATTAAGGCTG 1380  
DB 18390 TGATTAAGCATATATCTGTTTAACTGAGGCTCATTAAGCCTTGTGTTTAAATTAAGGCTG 18449  
QY 1381 TAAATTCGATTTCCAACTCTTACCAATTAATTAATGAAAGGCTGGAATGATGAACTGAC -TA 1439  
DB 18450 TAAATTCGATTTCCAACTCTTACCAATTAATTAATGAAAGGCTGGAATGATGAACTGAC 18509  
QY 1440 ATAAAGAAAAATTTGATGATGATTTATGCTATTCAGCAACCAATTAACCATCCAGAGC 1499  
DB 18510 ATAAAGAAAAATTTGATGATGATTTATGCTATTCAGCAACCAATTAACCATCCAGAGC 18569  
QY 1500 GACTTGGCAACCAATTTCTCAATTAATGATATGATGAAAGCAAGTTCTGATTTGCTCAAT 1559  
DB 18570 GACTTGGCAACCAATTTCTCAATTAATGATATGATGAAAGCAAGTTCTGATTTGCTCAAT 18629



QY 421 TCAACATCGTGAAGTGGAACTCCAAAGACGATGCTGTTGCTTGGCAGCTTCCGA 480  
|||||  
Db 480 TCAACATCGTGAAGTGGAACTCCAAAGACGATGCTGTTGCTTGGCAGCTTCCGA 539  
QY 481 AGAAGCTTACTACAGATGATGTTATCTTAACTGCTTGATATCAATAGAGATAC 540  
|||||  
Db 540 GGAAGCTTACACAGATGATGTTATCTTAACTGCTTGATATCAATAGAGATAC 599  
QY 541 TGTGATGCTTATATGCTTCTCATGAGATCATTAACATTACATTTCTTAAGAATGATT 600  
|||||  
Db 600 TGTGATGCTTATATGCTTCTCATGAGACATTAACATTACATTTCTTAAGAATGATT 659  
QY 601 ATGAGTAGAGATGCTGCTGTCAGAGAGCTTCCATCTGCTGTCGAGAAATCTGTCAA 660  
|||||  
Db 660 ATGAGTAGAGATGCTGCTGTCAGAGAGCTTCCATCTGCTGTCGAGAAATCTGTCAA 719  
QY 661 TTCAAGAACTTATCGCCGACAAAATAGCATTAACCTTCAAGAACAAATGCGGTACCTTC 720  
|||||  
Db 720 TTCAAGAACTTATCGCCGACAAAATAGCATTAACCTTCAAGAACAAATGCGGTACCTTC 779  
QY 721 TGTAGCAATTCAGGAATCAAAATTAATAACAAGAACAAAGCAACACTAATCA 780  
|||||  
Db 780 TGTAGCAATTCAGGAATCAAAATTAATAACAAGAACAAAGCAACACTAATCA 839  
QY 781 AGCAAGTCAAGTATGACATTTGATAGTCTTGAAGACAGCTCTACAACTGCTTGGAG 840  
|||||  
Db 840 AGCAAGTCAAGTATGACATTTGATAGTCTTGAAGACAGCTCTACAACTGCTTGGAG 899  
QY 841 TCAACGACATGTAGATCTGATGCGCTTGTCTTGTATTCAGACAAATCAACAGTGCAG 900  
|||||  
Db 900 TCAACGACATGTAGATCTGATGCGCTTGTCTTGTATTCAGACAAATCAACAGTGCAG 959  
QY 901 AGTGAAGTGTGACAGTGGCACAGGAGATCAATTAACACTTCACTCCCTTACCTCAAT 960  
|||||  
Db 960 CGCAATGAGTGTGCTGTACGACGAGGACCAATTAACCTTATTCCTTATTCACAACT 1019  
QY 961 GTCTGAATTTGGAAGAAAGATTCGCTGATTTATTCCTTGTATTCCTTCAACATTC 1020  
|||||  
Db 1020 GTCTGAATTTGGAAGAAAGATTTGCTGCTGATTTATTCCTTGTATTCCTTCAACATTC 1079  
QY 1021 GGTACGATTTCAAGGCGCAGACAGCAAGTCCAGCAACGACTCCGAGACCTAGTCCAG 1080  
|||||  
Db 1080 GGTACGATTTCAAGGCGCAGACAGCAAGTCCAGCAACGACTCCGAGACCTAGTCCAG 1139  
QY 1081 CCGGCAACCTGCGCAAACTTAAATAGACTCAA-----ATCTCTTGGT 1128  
|||||  
Db 1140 TCCGCAACCTGCGCAAACTTCAACAGCTCCAAAGCAATTCATTTGATAGAAATTTGGT 1199  
QY 1129 TAGTCAAGTGTACGAAAGATTTGGGAGAGATATGTATTCGAAGAAAGGCACTCTCTCG 1188  
|||||  
Db 1200 CAAAGAAAGCTGTTCGAAAGATAGGCGATGCTTATGCTTTGAGAGAGATGAGATTCCTCG 1259  
QY 1189 TTATGCTTTTGCAGAAAGATTTACCATCTGAAGCTTTAAAAATCTTGAAGCAAGTTATC 1248  
|||||  
Db 1260 TTATATTCAGGCAAGATTTTTCAGGAGAAAGAGCAGGAGATGATAGCAAACTGGC 1319  
QY 1249 AAAACAAGAGAGTGTTCACACTTTAACTGCTAAAAAGAAATTTCTCTCTCTCTGA 1308  
|||||  
Db 1320 CAAGCAGGAAGATTTATCTATAGCTAGAGCTTAAGAAAGTGAACCTCCATCTATGTA 1379  
QY 1309 CCAAGATTTTATGATTAAGCATATATCTGTTAACTAGAGGCTCAATAAAGCTTGTGTGN 1368  
|||||  
Db 1380 TAGAGATTTTAAATAGAGCTTATGACTTACTAGCAAGAAATTCATCAAGATTTACTTGA 1439  
QY 1369 AAATTAAGGCTGTAATTTGATTTCCAGGCTTATAGCAAAATTTATTAAGCGCTTGAATGA 1428  
|||||  
Db 1440 TAATTAAGGCTGTAATTTGATTTTGAAGCTTATGATTAAGCTTGAAGCAAGTCAAGGA 1499  
QY 1429 TGATGACATTAATTAAGAAATTTGATGATGATTTATTTGGCATTTCTTACACCAATTAAC 1488  
|||||  
Db 1500 TGCTCTAAGTATTAAGTCAAGTTAGTATGATTAATTTCTTGTAGCTCCGATTCG 1559

QY 1489 CCATCCAGACGACTTGGCAACCAATTTCTCAATTTAGTATATGAAAGCAAGTTGC 1548  
|||||  
Db 1560 TCATCCGAAGAGCTTTAGGAAAACCAATTTGCGCAATTTACTACTATATGATGATTTCA 1619  
QY 1549 TATTGCTCAATTTAGCTGATTAAGTATACACAGCTGAGTGTTAATTTTGAATGAATGA 1608  
|||||  
Db 1620 AGTAAAGAGTTGGCAGGCAAGTACACAGAAAGAGGTTATATCTTTGATCTCTGTA 1679  
QY 1609 TATATATGATGATGAAGGAATGCTATATGTAAGCCCTCATATGAGCCATATGCTAGAT 1668  
|||||  
Db 1680 TATATACAGTATGATGAGGAGTGTGATTAATCTCAATATGATGATGATGATGAT 1739  
QY 1669 TGGAAAGATAGCTTCTCTATTAAGGAAAGTTGACGCTCAAGCCATATCTAAAGAAA 1728  
|||||  
Db 1740 TAAAAAGATAGTTTGTCTGATGATGAGAGCGGAGCCAGCTTATGCTAAAGGAA 1799  
QY 1729 AGGTATCTACCTTCATCTCCAGACGAGATGTTAAAGCAATTCACACTGAGATAGTC 1788  
|||||  
Db 1800 AGGTTTACCCCTTCCTTCGACAGACATCAGATTCAGGAAATTAATGAGCAAAAGAGC 1859  
QY 1789 AGCAGCTATTTACAAATGCTGGAAGGGAAGAAAGCAATTCACCTGCTTGCATTCATA 1848  
|||||  
Db 1860 AGAAGCTATTCACAAAGCGGTGAAGGAGCTTAAGAAAGTCCACTGATGATGCTTGA 1919  
QY 1849 TATGCTGACATACAGTATGAGTTAAAAAGGTAATTTGATTAATTCCTATTAAGATCA 1908  
|||||  
Db 1920 CAATCTTCATATATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1979  
QY 1909 TTACCATTAATTAATTTTCTGCTGTTGATGATGATGATGATGATGATGATGATGATGAT 1968  
|||||  
Db 1980 TTACCATTAATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2039  
QY 1969 TACCTGGAAGATTTGTTTGCAGAGATTAAGTACTAGTGAACACCTTACAGACCTCC 2028  
|||||  
Db 2040 TAGCTTGAAGATCTTTTGGCAGCTGCAAGTACTATGATGATGATGATGATGATGATGAT 2099  
QY 2029 ACATCTTAATGATGATGAGGCAATGCGCAATGAGCATGCTGTTAGGCAAGAAAGCCAG 2088  
|||||  
Db 2100 GCATTCGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2159  
QY 2089 TGAAGATTCAAATTAAGACTTCAAGCGGATGAAGAGCCAGTGAAGAAACACTGCTGA 2148  
|||||  
Db 2160 AGCTGATACCAATTAAGCAAGGAAAGCAAGAGAGAGAAACCTGACAGCAAAACCTGA 2219  
QY 2149 GCCAGAGTCCCTCACTAGTGAAGA 2171  
|||||  
Db 2220 GGAAGAAACCCCTCGAGAGAGA 2242

RESULT 8  
AX343072 2523 bp DNA linear PAT 12-JAN-2002  
LOCUS AX343072  
DEFINITION Sequence 3 from Patent WO0198334.  
ACCESSION AX343072  
VERSION AX343072.1 GI:18152270  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
1  
AUTHORS Hamel,J., Ouellet,C., Charland,N., Martin,D. and Brodeur,B.  
TITLE Streptococcus antigens  
JOURNAL Patent: WO 0198334-A 3 27-DEC-2001;  
SHIRE BIOCHEM INC. (CA)  
FEATURES  
source  
1..2523  
Location/Qualifiers  
/organism="unidentified"  
/db\_xref="taxon:32644"  
/note="BVH-11"  
BASE COUNT 879 a 523 c 526 g 595 t  
ORIGIN  
Query Match 57.5%; Score 1374.2; DB 6; Length 2523;





QY 2126 ----- 2125  
Db 2220 GGAAGAAACCCCTCGAGAAGAAACCAAGGAGAAACAGAGTCTCCAAACCCAC 2279  
QY 2126 -----CGAGTAGAGAAACACCTGCTGAGCCAGAAAGTCCCTCAAGTAGACTGAAA 2178  
Db 2280 AGAGGAACCGAGAAAGATCCAGAGATCGAAGAACTCAGGCTGAGACTGAAA 2239  
QY 2179 AGTAGAAGCCCACTCAAGAGAGAGAGTTCCTTGCGAAATTAACGAGTCTAGCT 2238  
Db 2340 GGTTAGAGAAAAGCTGAGAGAGCTGAAGATTACTTGGAAGAAATCCAGATCAATTTAT 2399  
QY 2239 GAAAGCCATGCAACAGAACTAGCTGTTTACGAAATTTGACTCTTCAATTTAT 2298  
Db 2400 CAAGTCCAAATGCCAAGAGAGCTCACAGAGTAAAAAATATTACTATTGGCACCAC 2459  
QY 2299 GGATTAACATAGTATCATGCGAGAGAGAAATTAATCTTGCTGTTAAAAAGAGTAA 2358  
Db 2460 GGACAAACAATCTATTTATGCGAGAGCTGAAAACTATTGCTTTATTAAGAGAGTAA 2519

RESULT 9  
AX343073 2647 bp DNA linear PAT 12-JAN-2002  
LOCUS AX343073  
DEFINITION Sequence 4 from Patent M00198334.  
ACCESSION AX343073  
VERSION AX343073.1 GI:18152271  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified.

REFERENCE  
AUTHORS Hamel,J., Ouellet,C., Charland,N., Martin,D. and Brodeur,B.  
TITLE Streptococcus antigens  
JOURNAL Patent: WO 0195334-A 4 27-DEC-2001;  
SHIRE BIOCHEM INC. (CA)

FEATURES  
source 1..2647  
Location/Qualifiers  
/organism="unidentified"  
/db\_xref="taxon:32644"  
/note="BVH-11"

BASE COUNT 934 a 538 c 556 g 619 t  
ORIGIN

Query Match 57.5%; Score 1374.2; DB 6; Length 2647;  
Best Local Similarity 73.9%; Pred. No. 2.4e-283;  
Matches 1819; Conservative 0; Mismatches 539; Indels 102; Gaps 2;

QY 1 TTCTTACGAGTTGGACTGTATCAAGCTAGAACGTTAAGAAATATCGTTTCCTA 60  
Db 104 TGCCTTAGAAGTACTAGTTTGCATCAAGCTCAAACTGTAAAGAAATATCGTTTCCTA 163  
QY 61 TATAGATGAGAAACAGCGACGCAAAAAACGAGAGATTGACTCTGATGAGTTAGCAA 120  
Db 164 TATAGATGAGAAACAGCGACGCAAAAAACGAGAGATTGACTCTGATGAGTTAGCAA 223  
QY 121 GCGGAGAGATCAATGCTGAGCAATCGTCATCAAGATTAACCAACCAAGGCTATGTCAC 180  
Db 224 GCGGAGAGATCAACCGCCGACAAATCGTCATCAAGATTAACCAAGGCTATGTCAC 283  
QY 181 TTCACATGGGACCATATCATTTATCAATGTAAGTTCCCTATGACGTTATCATCAG 240  
Db 284 CTCGATGAGAGACCATATCATTTATCAATGTAAGTTCCCTATGATGTCATCATCAG 343  
QY 241 TGAAGATTTACTCATGAAAGATCCAACTATAGCTAAAGATGAGATATTTGTAATGA 300  
Db 344 TGAAGACTCTCTCATGAAAGATCCGAATTAATCAAGATTAACGATTTGTCATGA 403  
QY 301 GGTCAGAGGTGATATGTTATCAAGTATGATGAAATTAATTAATTAATTAATTAATTA 360  
Db 404 AATCAAGGAGGTGATATGTTATCAAGTATGATGAAATTAATTAATTAATTAATTAATTA 463

QY 361 TGCCACGCGGATTAACGTCCGTACAAAAAGAGAAATCAATCGACAAAAACAAGACATAG 420  
Db 464 AGCTCATGCGGATTAATGTCGTACAAAAAGAGAAATCAATCGGCAAAAAACAAGAACATAG 523  
QY 421 TCAACATGCTGAGAGTGAAGTCCAGAAAGAGATGTCGTTCCTTGCGACGTTGCA 480  
Db 524 TCAGCATGCTGAGAGAGAGGACTTCAGCAACGAGATGTCGTTCCTTGCGACGTTGCA 583  
QY 481 AGGAGCTATCTACAGATGATGATGATTAATGTTAATGCTTATCATATCATAGAGATAC 540  
Db 584 GGGAGCTACACACAGATGATGATTAATGTTAATGCTTATCATATCATAGAGATAC 643  
QY 541 TGTGATGCTTATATCGTTCCTCATGAGATCAATACATTAATTCCTTAAGATGAGTT 600  
Db 644 GGGGATGCTTATATCGTTCCTCATGAGATCAATACATTAATTCCTTAAGATGAGTT 703  
QY 601 ATCAGCTAGCGAGTTGGCTGCTGAGAAAGCCTTATCTGTCGAGAGAAATGTCAAA 660  
Db 704 ATCAGCTAGCGAGTTGGCTGCTGAGAAAGCCTTATCTGTCGAGAGAAATGTCAAA 763  
QY 661 TTCAAGAACTATGCGCCGACAAATAGAGATTAACACTTCAAGAACAACTGGTACCTTC 720  
Db 764 TTCAAGAACTATGCGCCGACAAATAGAGATTAACACTTCAAGAACAACTGGTACCTTC 823  
QY 721 TGTAGCAATCCAGAGAACTAACAATACTACACAAGCAACAAGCAACACTAACAGTCA 780  
Db 824 TGTAGCAATCCAGAGAACTAACAATACTACACAAGCAACAAGCAACACTAACAGTCA 883  
QY 781 AGCAAGTCAAGATATGATCATTTGATGCTTGAAGAGCTTCAAACTGCTTTGAG 840  
Db 884 AGCAAGTCAAGATATGATCATTTGATGCTTGAAGAGCTTCAAACTGCTTTGAG 943  
QY 841 TCAACGATCTGAATCTGATGCTGCTTGTGATTCACACAAATCAAACTGCAAC 900  
Db 944 TCAACGATCTGAATCTGATGCTGCTTGTGATTCACACAAATCAAACTGCAAC 1003  
QY 901 AGCTAGAGTGTGCTGAGTGCACACAGAGATCAATACACTTCACTCTCTCAAT 960  
Db 1004 CGCGAGAGTGTGCTGAGTGCACACAGAGATCAATACACTTCACTCTCTCAAT 1063  
QY 961 GTCTGAATTTGAGAGAGAAATGCTGATTAATTCCTGTTATGTTCAACATATG 1020  
Db 1064 GTCTGAATTTGAGAGAGAAATGCTGATTAATTCCTGTTATGTTCAACATATG 1123  
QY 1021 GGTACCAATCCAGAGCCGAGAACCAAGCCAAACCGGAACTGATCCAGG 1080  
Db 1124 GGTACCAATCCAGAGCCGAGAACCAAGCCAAACCGGAACTGATCCAGG 1183  
QY 1081 CCCGCAACCTGACCAAAATCTTAATAATAGACTCAA-----ATTCTCTTTGGT 1128  
Db 1184 TCGGCAACCTGACCAAAATCTTAATAATAGACTCAA-----ATTCTCTTTGGT 1243  
QY 1129 TAGTCACTGCTACGAAAGTTGGGAGAGATATGATTTGAGAGAAAGGCAATCTCG 1188  
Db 1244 CAAAGAAAGCTGTTGAAAGATGAGGATGTTATGCTTTGAGAGAAAGGATGTTCTCG 1303  
QY 1189 TTAGTCTTTGCGAAAGATTTACCATCTGAAATCTGTAATTAATCTGTAACCAATATC 1248  
Db 1304 TTAGTCTTTGCGAAAGATTTACCATCTGAAATCTGTAATTAATCTGTAACCAATATC 1363  
QY 1249 AAAACAAGAGAGTGTTCACACACTTAACGTCAAAAAAATAATGTTGCTCTCGGA 1308  
Db 1364 CAACAGAGAAAGTATATCTCAATAGCTAGAGGCTTAAGAAAACCTGCTCCATATGAGA 1423  
QY 1309 CCAAGAAATTTATGATAAAGCATATATCTGTTAATCTGAGGCTCATAAAGCTTTGTGN 1368  
Db 1424 TCGAGAAATTTTACATTAAGGCTTATGACTTACGAAAGATTCACCAAGATTTACTTGA 1483  
QY 1369 AAATTAAGGCTGTAATCTGTTTCCAAAGCTTAGACAAATTAATTAAGCTTGAATGA 1428  
Db 1484 TAATTAAGGCTGTAATCTGTTTCCAAAGCTTATGACTTACGAAAGATTTACTTGA 1533  
QY 1429 TGAATCGACTAATTAAGAAATAATGATGATTAATGTCATTCCTAGACCAATTAAC 1488



Db	1544	TGTCTCAAGTGAATAAGTAACTAGTATGATATTTCTTGGCTTCTTACGCTCCAGTTTC	1603
Qy	1489	CCATCCAGAGCGACTTGGCCAAACCAAAATTCGAAATTTGATATGATGAGAGCAAGTTTCG	1548
Db	1604	TCATCCAGAACGTTTAGGAAACCAAAATGCGCAAAATTACCTACACTGATGATGAGATTCA	1663
Qy	1549	TATGTCTCAATTAGCTGATATGATATCAACGTCAGATGGTTACATTTTGTGATGAACTGA	1608
Db	1664	ACTAGCCAAATTTGGCGAGCGCAAGTACCAACAGAAAGCGGTTATATCTTTGATCTCGTGA	1723
Qy	1609	TATTAACACTGATGAAAGAGATGCATATGTAAAGCGCTCATATGAGGCGCATAGTACAGTGA	1668
Db	1724	TATTAACACTGATGAGGGGGATGCTCATATTAATCTCCATATGACCATATAGCCACTGGAT	1783
Qy	1669	TGGAAAGATAGCCTTTCTGATATAGGAAAGTTGCGAGCTCAAGCCTATATCTAAAGAAAA	1728
Db	1784	TAAAAAGATAGTTGTCTGTAAGCTGAGAGAGCGGAGCGCCAGCGCTTATGCTAAAGAGA	1843
Qy	1729	AGGTATCTTACCTCCATCTCCAGAGCGAGATGTTAAAGCAAAATCCAACTGAGATATGTC	1788
Db	1844	AGGTTTGACCCCTCTTCCAGACACATCAGGATTCAGGAAATATCTGAGGCAAAAGGAC	1903
Qy	1789	ACGAGCTATTTCAATCGTGTGAAAGGGGAAAAACCAATTCACCTGCTGCACTTCATA	1848
Db	1904	AGAGCTATCTACAAACCGGTGAAAGCACTTAAGAAAGTGCCACTGATGCTGATGCTTGA	1963
Qy	1849	TATGTTTGAGCATACAGTTGAGGTTAAAAACGCTATTTGATTTATCTCATTAAGGATTA	1908
Db	1964	CAATCTTCAATATACGTGTAAGTCAAAAACGGTAGTTTAATCATATCCTCATATATGACCA	2023
Qy	1909	TTACCTAATATTAATTTGCTGGTTTATGATCATACATATCAAAAGCTCCAAATGCTTA	1968
Db	2024	TTACCAATCAATCAAAATTTTGAAGTGTGACCAAGGCTTTATGAGGCACTTAAGGGGCTA	2083
Qy	1969	TACCTTGAAGATTTGCTTGGCGACGTTTAAGTACTGATGAGAACACCCGACGAAGGTC	2028
Db	2084	TACTCTTGGAGATCTTTTGGCGACTGTCAAGTACTATGTCGAAACATCCAAAGCAAGCTCC	2143
Qy	2029	ACATTTCTAATGATGATGGGGCAATGCCAGTACATGTGTTAGGCAAGAAGACCAACAG	2088
Db	2144	GCATTCAGATTAATGGTTTTGGTAACGCTAGCGCATGTTCAAAGAAACAAAATATGCTCA	2203
Qy	2089	TGAAGATCCAAATTAAGACTTCAAGCGGATGAAG-----	2125
Db	2204	ACCTGATACCAATCAAAACGAAAAACCAAGCGAGGAACCTCAGACAGAAAAACCTGA	2263
Qy	2126	-----	2125
Db	2264	GGAGAAACCCCTCGAGAAAGAGAAACCAAGACGAGAAACCAAGTCTCCAAAAACCAAC	2323
Qy	2126	-----CCAGTAGAGAGAAACACTGCTGAGCCAGAGAGTCCCTCAAGTAGAGATGAAAA	2178
Db	2324	AGAGAGACCAAGAAAGAAATCAACAGAGAAATCAAGAAAGACCTCAGCTGAGAACTGAAAA	2383
Qy	2179	AGTAGAAGCCCACTCAAGAGAGACAGAACTTTGCTTGGGAAAGTAGACGATTTAGTCT	2238
Db	2384	GGTTGAGAAAAAATCGAGAGGCTGAGAGATTACTTGGAAAAAATCCAGATCCAAATTAT	2443
Qy	2239	GAAAGCCATGCAACAGAACTCTAGCTGTGTTAGGAATATTTGACTCTTCAATTAT	2298
Db	2444	CAAGTCCATGCGCAAGAGACACTCAGAGATTAATAAAATTAATTAATTTGAGCAACCA	2503
Qy	2299	GGATTAACATAGTATCATGCGAGAGACAGCAAAAAATTAATCTTGGGTTGTTAAAAAGAGTA	2358
Db	2504	GGACAAACAATATATATATATGCGCAAGAGCTGAAAAAACTATTGGCTTTATTTAAAGAGATGA	2563
RESULT 10			
LOCUS	AF1818954	2457 bp	DNA linear BCF 11-FEB-2001
DEFINITION	Streptococcus pneumoniae pneumococcal histidine triad protein B precursor (phbB) gene, partial cds.		

ACCESSION	AF18954
VERSION	AF18954.1
KEYWORDS	GI:12744741
SOURCE	
ORGANISM	
REFERENCE	Streptococcus pneumoniae.
AUTHORS	Bacteria: Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.
TITLE	1 (bases 1 to 2457)
JOURNAL	Adamou, J. E., Heinrichs, J. H., Erwin, A. L., Walsh, W., Gayle, T., Dormitzer, M., Dagan, R., Brewh, Y. A., Barron, P., Lathigra, R., Langemann, S., Koenig, S. and Johnson, S.
MEDLINE	Identification and characterization of a novel family of pneumococcal proteins that are protective against sepsis
PUBMED	Infect. Immun. 69 (2), 949-958 (2001)
REFERENCE	2 (bases 1 to 2457)
AUTHORS	Adamou, J. E., Heinrichs, J. H., Erwin, A. L., Walsh, W., Dormitzer, M. and Johnson, S.
TITLE	Direct Submission
JOURNAL	Submitted (03-NOV-2000) Molecular Microbiology, MedImmune, Inc., 35 West Watkins Mill Road, Gaithersburg, MD 20878, USA
FEATURES	Location/Qualifiers
SOURCE	1. 2457
gene	/organism="Streptococcus pneumoniae"
CDS	/serotype="4"
	/db_xref="taxon:1313"
	1. .>2457
	/gene="phtB"
	1. .>2457
	/gene="phtB"
	/note="phtB; protective pneumococcal antigen"
	/codon_start=1
	/transl_table=11
	/product="pneumococcal histidine triad protein B precursor"
	/protein_id="AAK06759.1"
	/db_xref="GI:12744742"
	/translation="MKINKRYLAGSVAVIALSVCSYELGRYAGQDKESNRVAYIDG DQGAKEMLTPDEVSKEKREGINAEOIYIKITDGYVSHGDHYHNGKVPDAIIE ELMKPDNYOLKDSIDYNEIKGVYLYLDAKHANIRKEIKRQKER SHNNSRADNAVAARAQGRVYTTDDGYTFASDIIEDTGDAYVPGDHYHYPKSEL SASLIAAEAVVMNGKOSRPSSSSYMANAPRLSENHMLTTPPTYHOGENISL LRELYAKPLSERHVESDGLIFDPAQITSRKAGVAVPHGNHIFIPEDNSELEKRIA RIIPDLRSHNHVPDSREBSPQPPPEPSPQAPASNPIDKLVKAEARKVGDGV FEENGYSRYTPAKDLSAETAAAGIDSKLAKOESLSHKLGTAKTDLPSDSREFYKAVDL LARHODLLNKGROVDEPFDALNDLRLKGVSSDKVYLVEDIAFLAPIRHPRKGLP NAOTYTHDEIOVAKIAGKTAEDGYFEDRDTISDGDAAVYVPMHTSHMTRKDSLS EAEFAAQAATAXEKGLTPSTDHODSSNTAKGAELIYNNKAAKVPDLDRMYNIQY TVEKNGSLIIPHYDAHNIKFEWDEGLIEARGYTLDELATVYKVEHPERPHS DNGGNASDHVORNRKNGQADTNTQTEKSEKSPOTERKEETPREEKPOSEKPSPPPT EENPESESEEPVETEKVEEKREADELLGIQDPIIKSNAKETLTGLKNMLLPGT QDNNTIAAEAKLLALLKESK"
sig_peptide	1. 87
misc_feature	/gene="phtB"
	247. .264
	/gene="phtB"
	/note="Region: histidine triad"
misc_feature	616. 633
	/gene="phtB"
	/note="Region: histidine triad"
misc_feature	718. .792
	/gene="phtB"
	/note="Region: coiled-coil domain"
misc_feature	940. .957
	/gene="phtB"
	/note="Region: histidine triad"
misc_feature	1396. .1488
	/gene="phtB"
	/note="Region: coiled-coil domain"
misc_feature	1657. .1674
	/gene="phtB"
	/note="Region: histidine triad"



QY	1978	AGATTGTTTGGCAGATTAAGTACTAGCTAGCAACCCCTGACGAAAGCTCCATCTTAA	2037
Db	1989	GGATCTTTTGGCGACTGTCACAGTACTATGTGCAACATCCAAACGAAAGCTCCCATTCAGA	2048
QY	2038	TGATGATGGGGCAATGCCAGTGACACTGTGTTAGGCAAGAAAGACACAGTGAAGATCC	2097
Db	2049	TAAATGTTTGGTAAACGCTAGCCAGCACCTGTTCAAGAAACAAAATGCTCAAGCTGATAC	2108
QY	2098	AAATAGAACTTCAAAAGCGGATGA-----	2121
Db	2109	CAATCAAAACGAAAAAACCAAGCGAGAGAAACCTCAGACAGAAAAAAGCTGAGAGAAAC	2168
QY	2122	-----AGAGCCAGT	2130
Db	2169	CCCTGAGAGAGAAACCGCAAAAGCAGAAACAGAGTCTCCAAAACCAACAGAGAGAAC	2228
QY	2131	AGAGAAACACTGCTGAGCCGAAAGTCCCTCAAGTAGAGACTGAAAAAGTGAAGCCCA	2190
Db	2229	AGAAATAATCCACGAGAGATTCAGAAAGAACCTTAGGTCGAGACTGAAAGGTTGAAGAAAAA	2288
QY	2191	ACTCAAAAGACAGAGAAGTTTGTCTTGCGAAATGAACGAGATTCAGTCAAAAGCCATGC	2250
Db	2289	ACTGAGAGAGGCTGAAGATTACTTTGGAAAAATCCAGATTCCAATTATCAATCCCATGC	2348
QY	2251	AACGAAACCTAGCTGGTTTACGAATPATTTGACTTCAAAATTATGATTAACATAG	2310
Db	2349	CAAAAGACACTCCACGAGATTAATAAATAATTTACTATTGTGCACCCAGACAAACAATAC	2408
QY	2311	TATCTGGCAGAGCAAGAAAAATTCCTTGCGCTGTGTTAAAGAGAGTAA	2358
Db	2409	TATTATGGCAGAGCTGAAAAACTTTGGCTTATTAAAGGAGAGTAA	2456

```

RESULT 11
LOCUS AF318955 2517 bp DNA linear BCT 11-FEB-2001
DEFINITION Streptococcus pneumoniae pneumococcal histidine triad protein D
ACCESSION AF318955
VERSION AF318955
KEYWORDS AF318955.1 GI:12744743
SOURCE
ORGANISM
.
Streptococcus pneumoniae.
Streptococcus pneumoniae
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE
1 (bases 1 to 2517)
Adamou,J.E., Heinrichs,J.H., Erwin,A.L., Walsh,W., Gayle,T.,
Dormitzer,M., Dagan,R., Brewar,Y.A., Barron,P., Lathigra,R.,
Langemann,S., Koenig,S. and Johnson,S.
Identification and characterization of a novel family of
pneumococcal proteins that are protective against sepsis
Infect. Immun. 69 (2), 949-958 (2001)
JOURNAL MEDLINE
21101045
PUBMED 1159990
2 (bases 1 to 2517)
Adamou,J.E., Heinrichs,J.H., Erwin,A.L., Walsh,W., Dormitzer,M. and
Johnson,S.
Direct Submission
Submitted (03-NOV-2000) Molecular Microbiology, MedImmune, Inc., 355
West Watkins Mill Road, Gaithersburg, MD 20878, USA
FEATURES
Source
location/Qualifiers
1..2517
/organism="Streptococcus pneumoniae"
/serotype="4"
/db_xref="taxon:1313"
1..>2517
1..>2517
/gene="phtD"
1..>2517
/gene="phtD"
/note="phtD: protective pneumococcal antigen"
/codon_start=1
/transl_table=11
/product="pneumococcal histidine triad protein D

```

[illegible]

Db	360	TGAATTCAGAGGCTGTTATATGTTTCAAGGTAGTGGAAATCTATGTTTACCTTAAAGGA	419
Oy	358	TGCTGCCACGCGGATTAACGTCGCTACAAAGAGAAATCAATGCACAAAACAGAGACA	417
Db	420	TGCAGCTCATCCGATTAATATTTGCGACAAAGAGAGATTAAGACGTCAAGAGCAGAGACA	479
Oy	418	TATGTCAACATCTGTGAGGTGGAACTCCAGAAACGATGGTGGTTCCTTGCCATTGCACGTTG	477
Db	480	CAGTCATAATTCACGGGGGTGGTTCT-----AACGATTCAGAGAGTATGTGCACCGCAGGC	533
Oy	478	GCAAGGAGCGATATCTPTACATATGATGGTTATATCTTATATGCTTCGTGATATCATAGAGA	537
Db	534	CAAAGGAGCGATATCAACGCGATATGGTTATATCTTCAATGATCATGTGATATCATTTAGAGA	593
Oy	538	TACTGTGATGCTTATATTCCTTCCTCATGAGATCATATACATATCATCTCCAAAGATGA	597
Db	594	CACGGGTGATGCTTATATTCGTTCTCTACGGCAGCATTTACATTCATTTCTTAAGATAGA	653
Oy	598	GTTATTCAGCTAGCGAGTTGGCTGCTGCAGAGCCCTTCATCTGCTGAGAGAAATCTGTC	657
Db	654	GTTATTCAGCTAGCGAGTTAGCTGCTGCAGAGCCT-----	688
Oy	658	AAATTCAGAACCTATGCGCGACAAAATAGCATTAACATCTTAACAAACAACTGGGTACC	717
Db	689	-----ATTGGAATTTGGAGAGGAGGATTCGCTCTTCTTCAAGCTTACTGTTATATATGC	740
Oy	718	TTCTGTAAAGCAATTCAGAGACTCAAAATGCTAACACAGCAACACAGCACTAACAAG	777
Db	741	AAATTCAGCTTACACAGATTTGTCAGAGAACCAATCTGATCTGCTACTCCAACTTATCA	800
Oy	778	TCAAGCAAGTCAAAATGATGACATTTGATAGTCTTGTGAACAGCTCTCAAACTGCTCTT	837
Db	801	TCA-----AAATCAAGGGGAAAACATTTCAAGCCTTTTAGTGAAATTTGTATGCTAAACCTT	857
Oy	838	GAGTCAACGACATGTAGATATCATATGCGCTTGTCTTTGATCCAGACAAATACAACTG	897
Db	858	ATCAGAAACGCGATGTGGAAATCATGTAGCCTTATTTTCGACCAGGCAAACTACAGACGTG	917
Oy	898	AACAGCTAGAGAGTTTTCATGTGCGCACAGAGATCATTTACACTTCATATCCCTTACTGCA	957
Db	918	AACCGCCAGAGGTGTAGCTGTCCCTCATGTGTAACATTTACACTTTATTCCTTTAGAACCA	977
Oy	958	AATGTCTGAATTTGGAAAGAACGAATTCGCTGATATTTCCCTCTGTTATCTGTTCAACCA	1017
Db	978	AATGTCTGAATTTGGAAAAAGAAATGCTGTGATTTCCCTCTGTTATCTGTTCAACCA	1037
Oy	1018	TTTGGTATCCAGATTTCAAGGCCAGACACCAACCAAGTCCACACCGACTCCGGAACCTTAGTCC	1077
Db	1038	TTTGGTATCCAGATTTCAAGACCAAGCAACCAACCAAGTCCACCAATGCACTCCGGAACCTTAGTCC	1097
Oy	1078	AGGCGCCGAACCTGCACCAATCTTTAAATATAGCTCAATTTCTC-----TTT	1125
Db	1098	AAGTTCGGAACCTGCACCAATCTCTCAACCGCTCCAAACCAATTCGAATGTAGAAATTT	1157
Oy	1126	GGTATGTACAGCTGTATACGAAGAATTTGGGAGAGATTTGATTTCGAAGAAAAGGCGATCTC	1185
Db	1158	GGTCAAGAAAGCTGTTCGAAAATATAGCCATGTTATGTCTTTTGAGGAAATGGAAGTTTC	1217
Oy	1186	TCGTTATATGCTTTTGCGAAGATTTTACATCTGAAATCTGTTAAAAAATCTTGAAGCAAGTT	1245
Db	1218	TCGTTATATATCCACAGCCAAAGATCTTTACAGAGAAACAGACAGGCAATGTATAGAAACT	1277
Oy	1246	ATCAAAAACAGAGAGTCTTCAACACATTTAACTGTTAAAAAAGAAAAATGTTGCTCTCG	1305
Db	1278	GGGCAACACAGAAAAGTTTATCTATAGCTTAGAGAGCTTAGAAAACCTGCATCTCCATCTAG	1337
Oy	1306	TGACCAAGAAATTTTATATGAATAAGCATATATATGTTAACTAGAGGCTCATAAAGCCTGTT	1365
Db	1338	TGATTCGAAATTTTACATATAGGCTTATGACTTACTAGCAGAAATTCACCAAGATTTACT	1397
Oy	1366	TGNAAAATTAAGGCTGTATATCTGATTTTCCAAAGCCTTAGCAAAATTTATGAACGCTTAA	1425

Db	1398	TGATATAAAGGTCACACAAGTGGATTTTGAGGCTTTGGATTAACCTGTTGGAAACGACTCAA	1457
Qy	1426	TGATGAATTCGATTAATAAAGAAAAATTTGGTAGATGATTATTTAGGCACTTCAGACCAAT	1485
Db	1458	GGATGCGCAAGTGAATTAAGTCAAGTTAGTGTATATTTCTGGCTCTTAAAGTCGAT	1517
Qy	1486	TACCATTCAGAGCGACATTTGGCAACCAAAATTCGAAATTAGTACTGAAGACGAAGT	1545
Db	1518	TGGATATCCGAGACGTTTAGGAAAAACCAAAATGCGCAAAATTACCTACACTGATGATGAAT	1577
Qy	1546	TGCTATTGCTCAATTAAGCGATAGATATTAACAACGTCAGATGGTTTCAATTTTTTGGATGACAA	1605
Db	1578	TCAAGTAGCGCAAGTTGGAGCGCAAGTATACACAGAAAGCGGTTATATCTTTGATCTTCG	1637
Qy	1606	TGATATTAATCAAGTGAAGAAGATGCATATATGACGCTCATATATGAGGCAATGACACTG	1665
Db	1638	TGATATTAACCATGATGAGGGGAGATGCTATCTATCTCCACATATGACCACATAGGCCACTG	1697
Qy	1666	GATTGGAAAAGATAGCCCTTTCTGATTAAGGAAAAAAGTTGCAGCTCAAGCCTATACTAAGA	1725
Db	1698	GATTAAAAAAGATGATTGTCTGTAAGCTGAGAGAGACGGCAGACGCGGTATGCTAAGA	1757
Qy	1726	AAAGGATTCCTACTCCATCTCCAGACGCGAGATGTTAAAGCAATCCAAGTGGAGATAG	1785
Db	1758	GAAAGGTTTGACCCCTCTTCGACAGACCATTCAGGATTCAGGAAATCTGAGGCAAAAAG	1817
Qy	1786	TCCACAGCATTTTACAAATCGTAGAAAGGGGAAAAAGCAATTCACACTCGTTGCACTCC	1845
Db	1818	AGCAGACGCTATCTACAAACCGGTAAAGCACCTTAAGAAGTGCACCTTGATCGATACC	1877
Qy	1846	ATATATGTTGAGCATAGATTGAGGTTAAAAACGGTAATTTGATTTATCTCATTAAGA	1905
Db	1878	TTACATATCTTCAATATCTAGTAGAAGTCAAAAACGGTATTAATCATACCTCATTAAGA	1937
Qy	1906	TCATTAACCATATATTTAAATTTGCTTGGTTGATGATACACATACAAAGCTCCAATAG	1965
Db	1938	CCATTATACCATCAATCAATTTGAATGGTTTGACCAAGCCCTTTATGAGGCAACCTAAGG	1997
Qy	1966	CTATACCTTGGAAAGATTTGTTGCGACGATTAAGTACTACGTAGAACAACCCGTGACGAG	2025
Db	1998	GTATATCTGTGAGATCTTTTGGCAGCTGTCAAGTACTATGTGCAACATCCAAAGCAAG	2057
Qy	2026	TCCACATTCCTAATGATGATGGGGCAATGCCAGTGAAGATGTTAGGCCAAGAAAGACCA	2085
Db	2058	TCCGATTCAGATATATGTTGTTGTAAGCGTACGACCATGTTCTTAATAATTAAGTGA	2117
Qy	2086	CAGTGAAGATCCAAATAGAACTTCAAAAGCGGATGAAGACCGAGTAAGGAAACACCTGC	2145
Db	2118	CCAAACAGATTAACCTGATGAAGATGAAGAACATGATGAAGTAACTGAGCCAACTACACC	2177
Qy	2146	TGAGCCAGAG 2156	
Db	2178	TGAATCTGATG 2188	
RESULT 12			
AE007403			
LOCUS			
DEFINITION	AE007403	11931 bp	DNA linear BCT 31-AUG-2001
ACCESSION	AE007403	AE005672	
VERSION			
KEYWORDS	AE007403.1	GI:14972469	
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
2 (bases 1 to 11931)  
Tettelin,H., Nelson,K.E., Paulsen,I.T., Eisen,J.A., Read,T.D., Peterson,S., Heidelberg,J., DeBoy,R.T., Haft,D.H., Dodson,R.J., Durkin,A.S., Gwinn,M., Kolonay,J.F., Nelson,W.C., Peterson,J.D., Umayam,L.A., White,O., Lewis,M.R., Radune,D., Holtzapfle,E., Khouiri,H., Wolf,A.M., Uteerback,T.R., Hansen,C.L., McDonald,L.A., Feldblyum,T.V., Angiuoli,S., Gesuwan,P., Hickey,E.K., Holt,I.E., Loftus,B.J., Ujval,M.L., Yang,F., Smith,H.O., Venter,J.C., Dougherty,B.A., Morrison,D.A., Hollingshead,S.K. and Fraser,C.M.  
Complete genome sequence of a virulent isolate of Streptococcus pneumoniae  
Science 293 (5529), 498-506 (2001)  
11463916  
2 (bases 1 to 11931)  
Tettelin,H., Nelson,K.E., Paulsen,I.T., Eisen,J.A., Read,T.D., Peterson,S., Heidelberg,J., DeBoy,R.T., Haft,D.H., Dodson,R.J., Durkin,A.S., Gwinn,M., Kolonay,J.F., Nelson,W.C., Peterson,J.D., Umayam,L.A., White,O., Lewis,M.R., Radune,D., Holtzapfle,E., Khouiri,H., Wolf,A.M., Uteerback,T.R., Hansen,C.L., McDonald,L.A., Feldblyum,T.V., Angiuoli,S., Gesuwan,P., Hickey,E.K., Holt,I.E., Loftus,B.J., Ujval,M.L., Yang,F., Smith,H.O., Venter,J.C., Dougherty,B.A., Morrison,D.A., Hollingshead,S.K. and Fraser,C.M.  
Direct Submission  
Submitted (29-JUN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA  
Location/Qualifiers  
1. 11931  
/organism="Streptococcus pneumoniae TIGR4"  
/strain="TIGR4"  
/db\_xref="taxon:170187"  
129. 476  
/gene="SP0995"  
655. 795  
/gene="SP0996"  
/note="This region contains a match to at least one other gene that is not full length, and is not the result of a sequencing artifact; similar to GP:5019553; identified by sequence similarity; putative; IS630-Spn1, transposase orf2, truncation"  
956. 1156  
/gene="SP0997"  
956. 1156  
/gene="SP0997"  
/gene="SP0997"  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAK75114.1"  
/db\_xref="GI:14972470"  
/translation="MVASASSTSTQAEQVDSKELRALSQELDRKALATVSDPK IDARKAVLDKAKPEDSALTE"  
1303. 1602  
/gene="SP0998"  
1303. 1602  
/gene="SP0998"  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAK75115.1"  
/db\_xref="GI:14972471"  
/translation="MTEENSVPAGYIVVSLALLGVIAFWLIRKKESEIOLSTELI KVLGOLDAERAKVLAQKQLDQETLDFVAKENGSATETRIKIVEELKAIIDDKK"  
1833. 2540  
/gene="SP0999"  
1833. 2540  
/gene="SP0999"  
/note="similar to GP:6714962; identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="cytochrome c-type biogenesis protein CcdA"  
/protein\_id="AAK75116.1"  
/db\_xref="GI:14972472"  
/translation="METIVFLISVFLAGVLSFSPSCIFPLPYAGILLDDQESAKSF SLFGKVLMSGLIRLCTIAGISLFTLIGAGVFGHILVANNFRYGMGAIITILGL HOMEFLIKLEVOKSEFFKRSDSNRWSAFLGITTFSGWTPCIQPVLSVLAAS

gene  
CDS  
2551. 3108  
/gene="SP1000"  
/gene="SP1000"  
/codon\_start=1  
/transl\_table=11  
/product="chlorodoxin family protein"  
/protein\_id="AAK75117.1"  
/db\_xref="GI:14972473"  
/translation="WKVMEAGLSLSLVLMACGEETKQTQAQOPKQOITYOQIA VGKAPDPTLSQMSQKEVLSDFGKRYKLFMAWCGPCPKSPMEIMELAKDRDF EILVIVAPGIGETVEQFPQWFOEQGKDPVLYDKATFFQYQIORSIPTEVLDS OGKIKIGFAGISNADAARKEMN"  
3482. 4873  
/gene="SP1001"  
3482. 4873  
/gene="SP1001"  
/note="identified by match to PFAM protein family HM PF00324"  
/codon\_start=1  
/transl\_table=11  
/product="amino acid permease family protein"  
/protein\_id="AAK75118.1"  
/db\_xref="GI:14972474"  
/translation="MNIERTKNVSLDKTEMRHKLMDLILIGICAMVGTFTGT AAATLAPALVTSISALCYGLSALPFAEASRPATGAYSTVLALIGEFPAMLVG MWTMEFMTAISGVAAGAAVFKGLISQYIAPQALNGTFNPGAGFVDDLPILVLY LVSIVLNLAKAALRPNISVILYLFSALELVIGINIKDKNSNRPYFGQITGA STGIMAGASLMFEGFLGFESISMAVDEKTPQKNIPRIVLSIVTLVAVLVLT GVVHSHLVNDAAVAFALRSGISMAVYSLVILTLVICISMITYALSMYISLAR DGLVPAARELTETSKIRKNTILTLGLASVAAGMPLASIAFNLICTLAYLIMLAY GLIRLRKKGMPKAGEFETPLVPLIISIIICSMQYNMTWIAFLVLSGSI VFTYGYKHSITEE"  
5016. 5933  
/gene="SP1002"  
5016. 5933  
/gene="SP1002"  
/note="identified by match to PFAM protein family HM PF01297"  
/codon\_start=1  
/transl\_table=11  
/product="adhesion lipoprotein"  
/protein\_id="AAK75119.1"  
/db\_xref="GI:14972475"  
/translation="MKRONFLVILSVFLICGAGCKESQGTGKMTVSPYIVAM VKESGDLNDVRAIQSSGHSFSPSANDIAITDADFTVHSHLTLSMAGSLDPIK KSKYVLEASGEKTLERYPGLEDVAGGVDEKTLDPHTWLDPKAGEEQIADKL SEVDSEHKEITYQKNAQAFIKKAQELTKFQPKFEKATQKFTVHTAFSTYAKRFGIN OLGIAGISPEQEPSPROLTEIOEFVKYKVTIFTESNASSKVAETLVKSTVGKLT LNPESDPNDKTYLENLENNMSILAEELK"  
5941. 8460  
/gene="SP1003"  
5941. 8460  
/gene="SP1003"  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAK75120.1"  
/db\_xref="GI:14972476"  
/translation="MKINKRYLAGSVAVLAVCSYEELGRHOAGOVKRESNRVYIDG DOAGOKAENLTPDEVSKEGINAEQIYIKITDQGYVSHGDHYHYNGKQVYDIISE EILMKDRNYQKSDIYNEIKGYIVKDYKYVYLLDAADADIRKEIKRKROKH SHNKGGSNDQAVVAARQGRYTTDDGTIFNASOIIDETGQAVIYPPGHDHITPKNE LSASELAAEAAYNMGKQSSRSPSSNNANPAVPHGNHVIPTPYHOQGENISS LRLRLYKPLSERVESGLIFDPAQITSRTARVAAPDHGMHFIYEQMSLEKRI ARIILPRKSNHWVDSRPBQSPSTPEPSPQAPNPQAPSPNIDEXLVKEAVR KVGQGYEEENGVSARIPAKDLSAETAGDSIKAKOESHSKIGAKKTDLPSSDRF YNKAYDLARIHQDLLDNKQVDFEALDNILEKAKVPSKAYLVYDDIATFLPIKH PERIGKPAQOITTYDDDEIQVAKLAGKTTDEGYIFDPDITSDGDVAVTPHMTSHM IKKDSLSBAERAAQAQAAKEKGLTPSPDIDGDSGNTAKGAEAITYNRYAKAKVPLDR







OY	838	GAGTCACGACATGTAGATGTGATGGCTCTGTTTGATATCCAGCAAAATTCACAAAGTC	897
D	201147	ATCAGAACGCATGTGGAAATGTGATGGCTTATTTTGACCCAGCCAAATTCACAAAGTC	201206
OY	898	AACAGCTGAGAGTGTGGCAGTGCACACGAGAGATCATTTACACTTCCTTACTCTCA	957
D	201207	AACGCCAGAGGTGAGTGTGCTTCATAGGTAACCATTTACACGTTTATCCCTTATGAA	201266
OY	958	AATGCTCTGAATTTGGAAGACGAATTCGCTCGATTTATTCCTTCGTTATCGTTCAAACCA	1017
D	201267	AATGCTCGAAATTTGGAAGAAACGAATTCGCTCGATTTATTCCTTCGTTATCGTTCAAACCA	201326
OY	1018	TTGGGTACCAATTCGAAGGCCGAGAACCAACGACGACCAACGACACCTCCGAACTAGTCC	1077
D	201327	TTGGGTACCAATTCGAAGAACCAACGACGACCAACGACACCTCCGAACTAGTCC	201386
OY	1078	AGCCCGCAGACCTCAGCAAAATCTTAAATAGACTCAA-----ATTCTCTTT	1125
D	201387	AAGTCCGCAACCTCGACACCAATTCCTCAACGACGCTCCAGCAATTCATATGAAATTT	201446
OY	1126	GGTTATGTCAGCTGTTAGCAAAAAGTTGGGGAGAGATATGTATTCGAAGAAAAGGCGATCTC	1185
D	201447	GGTCAAAAGAACCTGTTCGAAAAAGTAGCGATGCTTATGCTTTGAGAGGAATGAGATTTTC	201506
OY	1186	TCGTTATGCTTTTGCGAAAGATTACATCTGAAACGTGTAAAAATCTTGAAGCAAGTT	1245
D	201507	TCGTTATATTCACAGCAGAGATCTTTTACGACGAAGAACAGCAGCGCATTTATAGCAAACT	201566
OY	1246	ATCAAAACAGAGAGTGTTCACACACTTTAAGCTCTGAAAAAGAAAATTTCTCTCTCG	1305
D	201567	GGCCAGCAGAGAAAGTTTATCTCATAGCTAGAGCTTAAAGAAAAGTACCTCCCATCTAG	201626
OY	1306	TGACCAAGAAATTTTATGATTAAGCATATATCTGTATAGAGGCTCATTAAGCCTTGT	1365
D	201627	TGATCGAAATTTTACATAAGGCTTATGACTTACTGCAAGAAATTCACCAAGATTTACT	201686
OY	1366	TGNAATATGAGGTGTGTAATCTGATTTTCCAAAGCCTTAGACAATTTATTAAGCCTTGA	1425
D	201687	TGATATATTAAGGTGACACAGTTGATTTTGGAGCTTGTGGATACCTGTTGGAAGACTCA	201746
OY	1426	TGATGTAATCGCATATAAAGAAAAATTTGATAGATTTATTTGGCATCTCTACACCAAT	1485
D	201747	GGATGTCTCAGTATATAAGTCAAGTTATAGTGAAGATATCTTGCCCTTCTTACTCCGAT	201806
OY	1486	TACCCATCCAGAGGACTTGGCAAACTCTCAATTTGAGTATATCGAAGACGAAGT	1545
D	201807	TCGTATCTCAAGAGTTTAGAAGAAACCAATAGCCAAATTAACCTACACTGATGATGAT	201866
OY	1546	TCGTATTTCTCAATTTAGCTGATAGTATATCAAGCTCAGATGGTTTACTTTTATGAACA	1605
D	201867	TCAAGTACCAAGTTTGGGAGGCAAGTACACAAAGAAAGGGTTATATCTTTATCTCTCG	201926
OY	1606	TGATATATTCAGTATGATGAAGAGATGATATGTAAACCTCATATGGGCCATAGTCACTG	1665
D	201927	TGATATATTCAGTATGATGAAGAGATGATATGTAAACCTCATATGGGCCATAGTCACTG	201986
OY	1666	GATTGAAAAAGATAGCCTTTCTGATTAAGGAAAAAGTTGCAGCTCAAGCCTATACTAAAG	1725
D	201987	GATTGAAAAAGATAGTGTGTGTAAGCTGAAAGAGCGGACCCAGCCTATGCTTAAAG	202046
OY	1726	AAAAGGTATCTTACTCTCATCTCCAGACGACGATGTAAAGCAAAATTCAACTGGAAGTAG	1785
D	202047	GAAAGGTTTGAACCCCTCTTCGACAGCATCAGAGATTTCAGGAAATATCTGAGCCAAAAG	202106
OY	1786	TGCAGCAGCCTTTTACATCGTGTGAAGGGGAAAAAGCAATTCCTACTGTTGACATTC	1845
D	202107	AGCAGAACTTATCTACACCGCGGTGAAAGCAAGCTTAAGAAAGTCCACTTGTATGCGATGCC	202166
OY	1846	ATATATGGTTGAGCATACAGTGTGAGTTAAAGAAAGGATTTGATTAATTCCTCATAGA	1905
D	202167	TTACATATTCATATATACGTAGAAAGTCAAAAACGGATGTTTAAATCATACTCATTAATGA	202226
OY	1906	TCATATTCCATATATTTAAATTTGCTTGTTGATGATCACACATATCAAGCTCCAAATGG	1965

Db	202227	CCATTACCATCATCAAAATTTTGAGTGGTTTACGCAAGGCCCTTTATGAGGCCTTAAGG	202286
Qy	1966	CTATACCTTTGGAAGATTTTGGTTTGCAGACGTTTAAGTACTACGTAGAACAACCCCTGACGAAAG	2025
Db	202287	GTATACCTGTTGAGATCTCTTTGGCGACGTGTCAAGTACTATGTGCAACATCCAAACGAAAG	202346
Qy	2026	TCCACATTTCTATGATGATGGGCGCAATGCCAGTGAACATGTGTGTTAAGCAAGAAAGACCA	2085
Db	202347	TCCCGATTTCAGATTAATGTTTGGTAACCGCTAGCCGACATGTTCAAAGAAACAAAATG	202406
Qy	2086	CAGTACAGATCCAAATTAAGAACTTCAAAAGCGATGAAAG	2125
Db	202407	TCAACTGTATACCAATCAACGGAAGAAAACCAACGAGGAGAAACCTCAGACAGAAAAACC	202466
Qy	2126	-----	2125
Db	202467	TGAGGAGAAACCCCTCGAGAGAGAAACCGCAAGCGAAACCAAGAGTCTCCAAAACC	202526
Qy	2126	-----CCAGTACGAGAGAAACACACCTGCTATAGCCAGCAAGATGCCCTCAAGTGAAGACTGA	2175
Db	202527	AACAGACGAAACCGAAGAAAGATCACACAGAGGAATTCAGAAAGAACTCAGCTGAGAGCTGA	202586
Qy	2176	AAAAATAGAACCCCACTCACTCAAAAGACAGAAATTTTGGTTCGGAAGATTAACGATCTCTAG	2235
Db	202587	AAAGTTGGAAGAAAAATCTGAGAGAGGCTGSAATTTACTTGGAAAAATTCAGAGATTCAT	202646
Qy	2236	TCTGAAAGCCATGCAACAGAAAGCTCTAGCTGTTTACGAATTAATTTGACTCTTCAT	2295
Db	202647	TATCAAGTCCAAATGCGCAAAAGACACTCTCACAGATTTAAAAATTTCTATTGTTGGCAC	202706
Qy	2296	TATGATTAACATAGTATCATGATGGACGAAGACAGAAATTAATCTTGCGTTTAAAGGAAG	2355
Db	202707	CCAGACAAACATACTATATATATGCGAGAAAGCTAAAAAATCTATTGGCTTTATTAAAGSAGAG	202766
Qy	2356	TAA 2358	
Db	202767	TAA 202769	
RESULT 14			
LOCUS	ARI120270	2290 bp	DNA linear PAT 16-MAY-2001
DEFINITION	Sequence 65 from patent US 6159469.		
ACCESSION	ARI120270		
VERSION	ARI120270.1 GI:14103846		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 2290)		
AUTHORS	Choi,G.H., Kunsch,C.A., Barash,S.C., Dillon,P.J., Dougherty,B., Fannon,M.R. and Rosen,C.A.		
TITLE	Streptococcus pneumoniae antigens and vaccines		
JOURNAL	Patent: US 6159469-A 65 12-DEC-2000;		
FEATURES	Location/Qualifiers		
source	1..2290		
BASE COUNT	766 a 474 c 498 g 547 t	5 others	
ORIGIN			
Query Match	41.3%; Score 987.6; DB 6; Length 2290;		
Best Local Similarity	67.7%; Pred. No. 1e-200;		
Matches 1481; Conservative 0; Mismatches 645; Indels 60; Gaps 5;			
Qy	1	TTCTTAGCAGTGGAGCTGTATCAAGCTAGAAAGGTTAAGAAAA--TAAATCGTGTTC	57
Db	4	TTCTCATGAACTTGGTGTGTCACCAAGCTGTCTGAGTTAAGAAAGAGTCTAAATCGAGTTTC	63
Qy	58	CTATATAGTGTGAACAAGACGCAAAAGAAACGAGAAATTTGACTCCGATGAGAGTTAG	117
Db	64	TTATATAGATGGTGTATCGAGGCTGCTCAAAAGGCAAAATCTTGACACACAGATGAAGTCAG	123

OY	118	CAGGGTGAAGGANTCAATGCTGTAGGCAAAATGCTATCAAGATAACAGACCAGGCTATGT	177
OY	118		177
Db	124	TAAAGGGAGGGGGATCAACCGCGCAACAAATGNNATCAAAATTAACGATTAACGATTAATGT	183
OY	178	CACCTTCACATGGCGCCACTATCTATTATTTCAATGGTAAGGTCTCTTATGACGCTATCAT	237
Db	184	GACCTCTCATGTGGAGACCTTTATCTTTACTATTAATGGCAAGGTTCTTTATGATGCCATCAT	243
OY	238	CAGTGAAGAATTACTCATGAAGAATCCAAACTATTAAGCTTAAAGATGAGGATATTGTTAA	297
Db	244	CAGTGAAGAAGCTCTCATGAAGAATCCGAATTTATAGGTTGAAGATTCAGACATGTCA	303
OY	298	TGAGGTCAAGGTTGATATGTATTCAAGTATGATGAAATTAATATGTTTACTTTAAGA	357
Db	304	TGAATTCAGAGGGTGTATATGTTCATTAAGGTAAACGGTAATTAATGTTATACCTTTAAGA	363
OY	358	TGCTCCCAACCCGATTAACGTCCTACAAAAGAGAAATCAATCGACAAAAAACAAGAGA	417
Db	364	TGCAGCTCATGTGGGATTAATTTGGCAAAAAGAGAGATTAAACGTTCAGACAGAGACG	423
OY	418	TAGTCACATCGTGAAGGTGGAACCTCCAGAAAAGATGGTGCTGTGCTTGGCAAGTTC	477
Db	424	CAGTCATTAATCAT-----AACTCAAGACGAGATTAATGCTGTGTCAGCGCAGAC	474
OY	478	GCAAGAGCGCTATCTACAGATATGTTTATATCTTAATAGCTTCGTGATATCATAGAGA	537
Db	475	CCAAAGAGCGTTATCAACGGATGTATGGGTATATCTTCAAAATGATTCGATATCATGTAGGA	534
OY	538	TACTGTGATGCTTATATATCGTCTCTCAAGAGATCAATTACATTAATTCCTTAAGAATA	597
Db	535	CACGGGTATCTTATATCGTCTCTCAAGCGCACCATTAACATTAATTCCTTAAGAATA	594
OY	598	GTTATCACTAGCGAGTTGGCTGCTGCAGAAAGCTTCTATCTGTCGAGAAATCTGTC	657
Db	595	GTTATCACTAGCGAGTTAGCTGCTGCAGAAAGCT-----	629
OY	658	AAATTCAGAAGACTATGCGCGACAANAATACGATTAACACTTCAAGACAACACTGGGTACC	717
Db	630	-----ATTGGAATGGGAAGGAGGAGATCGTCTCTTCAAGTTCTAAGTTAATAATAC	681
OY	718	TTCTGTAGCAATACAGAGACTACAAATTAATAACACAGAACAAACAGCAACTAACAAG	777
Db	662	AAATCAGACTCAACCAAGATTTGTACAGAAACCAATGTGACGTACATCCACATTTATCA	741
OY	778	TCAAAGCAAGTCAAAATAATGACATGTATGTCCTTTGAAAACAGCTCTACAAAATGCCCTT	837
Db	742	TCA---AAATCAAGGGAAAAAATTTCAACCTTTACGTAATTTGATGTAACCCCT	798
OY	838	GAGTCACAGAAATATAGAAATCTATAGCCCTTGTCTTGTATCCAGACAAATACAAAGTCG	897
Db	799	ATCAAGAACGCCATGTGAAATCTGATGGCTTATTTTGCAGCCAGGCCAAATACAAAGTCG	858
OY	898	AACAGCTAGAGGTGTGACAGTGCACACAGAGATCAATTACCACTATATCCCTTACTCTCA	957
Db	839	AACGCCAGAGGTTAGCTGTCTCTCATGTAAACATTAACCACTTATATCCTTATGAACA	918
OY	958	AATGTCTGAATTTGGAAGAAGAAATCGCTCTATTAATTCCTTCCCTTCTTATCGTTCAACA	1017
Db	919	AATGTCTGAATTTGGAAGAAGAAATGTGCTATTAATTCCTTCCCTTCTGTTATCGTTCAACA	978
OY	1018	TTGGGTACAGATTCACAGGCCAGAACACCAAGTCCACACACGACTCGGGAACCTAGTCC	1077
Db	979	TTGGGTACCAATTTCAAGACCAGAACCAACCAATGCCACATCGACTCGGAACCTAGTCC	1038
OY	1078	AGGCCCGCAACCTCCACCAAAATCTTAAATATAGACTCAAAATCTTC-----TTT	1123
Db	1039	AAGTCCGCAACCTCCACCAAAATCTTCAACCAAGCTCCACCAACATTCCAATTGATGAGAAAT	1098
OY	1126	GGTATAGTACAGCTGTGTAGAAAGTTGGGGAAGATATGATTCGAAAGAAAGGCATATC	1185
Db	1099	GGTCAAAAGAAAGCTTTGCAAAAAGTAGGCATGTTTATGTCTTTGAGAGAGATGGAGTTTC	1158
OY	1186	TGCTTATGTCTTTGCGAAAGATTTACACATCTGAACATGTTAAAAAATCTTGAAGCAAGT	1245

[illegible]

RESULT	15		
AX343074			
LOCUS	AX343074	2639 bp	DNA
DEFINITION	Sequence 5 from Patent WO019834.		linear
ACCESSION	AX343074		PAT 12-JAN-2002

VERSION	AX343074.1	GI:18152272
KEYWORDS	unidentified.	
SOURCE	unidentified	
ORGANISM	unidentified	
REFERENCE	1	
AUTHORS	Hamel, J., Ouellet, C., Charland, N., Martin, D. and Brodeur, B.	
TITLE	Streptococcus antigens	
JOURNAL	Patent: WO 0198334-A 5 27-DEC-2001;	
FEATURES	SHIRE BIOCHEM INC. (CA)	
source	Location/Qualifiers	
	1..2639	
	/organism="unidentified"	
	/db_xref="taxon:32644"	
	/note="BVH-11-2"	
BASE COUNT	889 a 518 c 567 g 665 t	
ORIGIN		
Query Match	41.1%;	Score 980.8; DB 6; Length 2639;
Best Local Similarity	67.8%;	Pred. No. 3e-199;
Matches 1473; Conservative	0; Mismatches 638; Indels 60; Gaps 5	
1	TTCTTACGAGTTGGGACCTGATCATCAAGCTTGAAAGCGTTAAGAAA---TAACTGGTTTC	57
173	TTTCTATGACTTGGTGGTCGACCAAGCTGTGACGTTAAGAAAGCTTAATGAGTTTC	232
58	CTATATAGATGAAAAACAAGCGACGCAAAAAAGAGAAATTGACTCTCATGAGGTTAG	117
233	TATATAGATGGTGTACAGCTGGTCCAAAGGACAAATTTGACACCAAGATGAAGTCAG	292
118	CAAGCTGAGGAATCATGCTGAGCAATTCGTATCAAGATPAACAGCAAGCTATGT	177
293	TAAAGAGAGGGGATCAACGCCAACAATTTGTTATCAAGATTACGATCAAGGTTATGT	352
178	CACCTCACATGGCGACCATATCATATATACAAAGTTAGGTGCTTACAGCTATCAT	237
353	GACCTCTAGGAGACCATATATCATATATATGCAAGTTCTTATGATGCAATCAT	412
238	CAGTGAAGATTTACATGATGAAGATCCAAACATTAAGCTAAGAAAGATGAGATATTGTTAA	297
413	CAGTGAAGAACTTCATCAATAAAGATCCGAATTTATCAGTTGAAGATTCAGACATTGTCAA	472
298	TGAGGTCAAGGTTGATATGTTATCAAGGTAGATGGAANAATCTATGTTACCTTAAGA	357
473	TGAATTCAGAGGTGGCTATGATGATTAAAGTAAAGCAAAATCTATGTTACCTTAAGA	532
358	TGCTTCCCGCCGCGGATACGTCCTGTCACAAAAGAAATCATCGACAAAACAGAGCA	417
533	TCCGCCCTCAAGTGGGCAAAATTTCCGCAAAAAGAGATTAAACGTCACAAACAGGAA	592
418	TAGTCAACATCTGTAAGGTGAATCTCAAGAACGATGGTGGCTGTGCTTGCGACGTTCC	477
593	CAGTCAATATCAT-----AACTCAAGACAGATTAATGCTGTGGCTGACGCCAGAC	643
478	GCAAGACGCTTACTACAGATGATGGTTATATCTTTAATGCTTCGATATCATAGAGA	537
644	CCAAAGACGTTATACAAAGCATGATGGGTATATCTTCATGATGATCATGATATCATGAGA	703
538	TACTGTGATGCTTTATATCGTTCCTCATGAGAAATCATTTACCATTCATTCTTAAGATA	597
704	CACGGGTATGTTTATATCTGTTCCACGCGCACCATTTACATTCATTCCTCAAGATTA	763
598	GTTATCAGTACGATGGTGGCTGTCGACAAACCTTCATCTGGTGGAGAAATCTGTC	657
764	GTTATCAGTACGATTTAGCTGTGTCGACAAAGCT-----	798
658	AAATTCAGAAACCTTATCGCCGACAAATATGCGATTAACATTCACAGAACAACTGGGTACC	717
799	-----ATTGGAATGGGAAGCAGGAGATCTGCTCTCTTCAAGTTCTAGTTATATATGC	850
718	TTCTGTAGCAATCCAGGAACCTACAAATTTCTAACACAAAGCAACAGACACATPAACAG	777
851	AAATCCAGTTCAACAGATTTGTGAGAGAACCAACATCTGACTGTCACTCCAACTTATACA	910

QY	778	TCACGCAAGTCACAAAGTAAATGACATTTGATTCCTTGAAACAGCCTCTACAAACGCTTT	837
Db	911	TCA-----AATCAAGGGGAAAAACATTTCAACGCTTTTACGTAATTTGATGTAAACCTT	967
QY	838	CAGTCACAGACATGTAGAAATCTGATGGCCCTTGCTTTGATCCAGACAAATACAAATCG	897
Db	968	ATCACAGAGCCATTAAGAAATCTGATGGCCCTTATTTTTCGACCGCAAGCAATATCAAGTCG	1027
QY	898	AACACCTGAGAGTGTTCGACGACACGGAGATCAATTACCACTTATCCCTTACTGTCA	957
Db	1028	AACCGCCAGAGGTGAGCTGTCCCTCAATGTAACCTTACCACTTATCCCTTATGAACA	1087
QY	958	AATGTCGAATTTGGAAGAGAAATCGCTGATATTTTCCCTTCGTTATCGTTTCAACCA	1017
Db	1088	AATGTCGAATTTGGAAGAAAGAAATTTGCTGATATTTTCCCTTCGTTATCGTTTCAACCA	1147
QY	1018	TTGGGTACCAATTCGAAGCCAGAAACCAACCAAGTCCACAACCGACTCGGAACTTAATCC	1077
Db	1148	TTGGGTACCAATTCGAAGCCAGAAACCAACCAAGTCCACAATCGACTCGGAACTTAATCC	1207
QY	1078	AGGCGCCGACCTCGACCAATCTTAAATAGCTCAAAATTCCTC-----TTT	1125
Db	1208	AAGTCGCAACCTCGACCAATCTTCAACAGCTCCACAGCAATCCAAATGATGAGAAAT	1267
QY	1126	GTTAGTCAGCTGCTAGCAAAAGTTGGGAGAGATATGATTTGAAAGAAAGGCGATCC	1185
Db	1268	GSTCAAGAACCTGTGTGAAAGATGAGCGATGGTTATGCTTTTGAGAGAAATGAGATTC	1327
QY	1186	TCGTTATGCTCTTCCGAAGATTTACCATCTGAACGTGTTAAATCTTTGAAAGCAAGTT	1245
Db	1328	TCGTTATGCTCTTCCGACCAAGATCTTTCGACGAAACAGACAGAGCGATGATGAAACT	1387
QY	1246	ATCAAAACAAGAGGTTCACACACTTAACTGCTAAAGAAAGAAATGTTGCTCTCG	1305
Db	1388	GGCCAAGCAGAAAGTTATCTCTAAGCTTAGAGGATAGAAAACTGACCTCCCATCTAG	1447
QY	1306	TGACCAACAATTTTATATATAAGCATTAATCTGTTAACTGAGGCTCATTAAGCCCTGTT	1365
Db	1448	TGATGAGAAATTTTACATATAAGCTTATGACTTACTAGCAGAAATTCACCAAGATTTACT	1507
QY	1366	TGNAATATAAGGCGTAAATTCGATTTCCAAAGCCTTAGACAAATTAATAGAACGCTTAA	1425
Db	1508	TGATATATAAAGTGTGACAAAGTGTATTTGAGTTTGGATTAACCTGTGGAACGACTCA	1567
QY	1426	TGATGAAATCGACTAATAAGAAAAATTTGGTAGATTTATTTGGCATTTCTTGACACCAAT	1485
Db	1568	GGATGCTCAAGTATTAAGTCAAAGTAGAGATGATATTCCTGCTTTTGCTCCGAT	1627
QY	1486	TACCATTCAGAGGACTTGGCAAAACCAATTCCTCAATTTGAGTATATCAAGAGCAAGT	1545
Db	1628	TCGTCATTCAGAACGTTTAGGAAAACCAATTCGCAATTTACTCTACATGATGATGAAT	1687
QY	1546	TCGATATTCGCAATTTAGCTGATAGTATATACACGTCAGATGGTATCATTTTGTATGAACA	1605
Db	1688	TCAGTATACCAAGTTTGGCAGGCAAGTACACAGAAAGAGGTTATATCTTTGATCCCTG	1747
QY	1606	TGATATATTCAGTATAGAGAGATGATATTTGAACGCTCATATGGCCATATGTCACCTG	1665
Db	1748	TGATATATACAGTATAGAGGGGATGCTATGTATACCTCCACATATAGCCATAGCCACTG	1807
QY	1666	GATTGGAANAATATGCGCTTTCTGTGTAAGAAAAAGTTGCAAGCCCAAGCCTATATCAAGA	1725
Db	1808	GATTGGAANAATATGCTTTGTGTGAGCTGAGAAAGGCGAGCCAGCCTTATGTCTTAAGA	1867
QY	1726	AAAAAGTATCTTACCTCATCTCCAGACGAGATGTTTAAAGCAAAATCCAACTGAGATAG	1785
Db	1868	GAAAGGTTTGACCCCTCTTCGACAGACCAACGAGATTGAGGAANAATCTGAGGCAAAAGG	1927
QY	1786	TGCAGACGCTATTTCACAATGCTGTGAAGGGGAAAAAGCAATTCCACTGTTGCACTTC	1845
Db	1928	AGCAGAAGCTATATCAACACGCGTGAAAGAGCTTAAGAAGGTCGCACTTGATCTGATGCC	1987



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 16, 2002, 23:10:35 ; Search time 340 Seconds

(without alignments)  
15823.600 Million cell updates/sec

Title: US-08-961-083-55

Perfect score: 2389  
Sequence: 1 TTCTTACGAGTTGGACGTCT.....TAAGTAGGAGAAAAATAAAC 2389

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_101002: \*  
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: \*  
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: \*  
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: \*  
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: \*  
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: \*  
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: \*  
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: \*  
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: \*  
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: \*  
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: \*  
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: \*  
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: \*  
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: \*  
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: \*  
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: \*  
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: \*  
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: \*  
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: \*  
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: \*  
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: \*  
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: \*  
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: \*  
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: \*  
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2388	100.0	2389	19	AAV27351	Streptococcus pneu
2	2388	100.0	2389	24	AB084819	S. pneumoniae SP03
3	2388	100.0	2451	21	AAA47604	Recombinant varian
4	2388	100.0	8195	19	AAV52227	Streptococcus pneu
5	1374.2	57.5	2523	21	AAA65731	Streptococcus pneu
6	1374.2	57.5	2647	21	AAA65736	Streptococcus pneu
7	1374.2	57.5	2647	24	ABK15103	DNA encoding Strep
8	1011.8	42.0	2478	21	AAA08557	S. pneumoniae 92 k
9	1003.8	42.0	2481	21	AAA05417	Streptococcus pneu

10	991	41.5	2531	21	AAA47605	Recombinant varian
11	990	41.4	2531	21	AAA47602	Recombinant varian
12	987.6	41.3	2290	19	AAV27356	Streptococcus pneu
13	987.6	41.3	2290	24	AB084824	S. pneumoniae SP04
14	980.8	41.1	2639	21	AAA65737	Streptococcus pneu
15	980.8	41.1	2639	24	ABK15104	DNA encoding Strep
16	869.4	36.4	2163	20	AAV52394	Streptococcus pneu
17	653.8	27.4	2359	19	AAV52376	Streptococcus pneu
18	548.4	23.0	3171	21	AAA65739	Streptococcus pneu
19	385.4	16.1	1342	19	AAV27414	Streptococcus pneu
20	385.4	16.1	1342	24	AB084882	S. pneumoniae SP10
21	385.4	16.1	1455	21	AAA65733	Streptococcus pneu
22	385.4	16.1	1455	21	AAA47603	Recombinant varian
23	385.4	16.1	1455	21	AAA05473	Streptococcus pneu
24	385.4	16.1	1455	21	AAZ91804	Streptococcus pneu
25	385.4	16.1	3120	21	AAA65730	Streptococcus pneu
26	385.4	16.1	5048	21	AAA65735	Streptococcus pneu
27	385.4	16.1	5048	24	ABK15101	DNA encoding Strep
28	385.4	16.1	6867	19	AAV52325	Streptococcus pneu
29	381.4	16.0	973	19	AAV52488	Streptococcus pneu
30	347	14.5	2528	21	AAA65738	Streptococcus pneu
31	347	14.5	2528	24	ABK15105	DNA encoding Strep
32	247.4	10.4	504	21	AAA08556	S. pneumoniae 20 k
33	243.4	10.2	1684	19	AAV52391	Streptococcus pneu
34	214.6	9.0	2475	24	ABN66838	Streptococcus poly
35	214.6	9.0	2478	22	AA500036	Streptococcus pyog
36	211.4	8.8	1146	21	AAA05814	Group B Streptococ
37	211.4	8.8	2466	24	ABN69535	Streptococcus poly
38	211.4	8.8	2466	24	ABN70334	Streptococcus poly
39	211.4	8.8	2469	21	AAA05811	Group B Streptococ
40	211.4	8.8	2469	22	AA500038	Streptococcus agal
41	209.8	8.8	2469	21	AAA65740	Streptococcus pneu
42	209.8	8.8	5215	20	AAK91105	Group B Streptococ
43	206.6	8.6	2472	21	AAA65741	Streptococcus pneu
44	201.6	8.4	1398	24	ABN66839	Streptococcus poly
45	201.2	8.4	819	24	ABK15106	DNA encoding Strep

#### ALIGNMENTS

RESULT 1  
ID AAV27351 standard; DNA; 2389 BP.  
AC AAV27351:  
DT 02-OCT-1998 (first entry)  
DE Streptococcus pneumoniae SP0036 nucleotide.  
XX  
KW Streptococcus pneumoniae: antigen; vaccine; infection; diagnosis;  
KW detection; pneumonia; otitis media; meningitis; ss.  
XX  
OS Streptococcus pneumoniae.  
XX  
FH Key Location/Qualifiers  
FT CDS 2..2389  
FT /tag= a  
FT /product= "SP0036"  
FT /transl\_except= (pos:1367..1369,aa:Xaa)  
FT /note= "no stop codon given; Xaa is unspecified"  
PN W09818930-A2.  
XX  
PD 07-MAY-1998.  
XX  
PF 30-OCT-1997; 97WO-US19422.  
XX  
PR 31-OCT-1996; 96US-0029960.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.

PI Choi GH, Hromocky J A, Johnson LS, Kunsch CA;  
XX MPI: 1998-272224/24.  
DR P-PSDB; AAM55090.  
XX  
XX Nucleic acid encoding antigenic peptide(s) from Streptococcus  
PT pneumoniae - or their epitope-containing fragments, useful in  
PS protective or therapeutic vaccines, and for diagnosis  
PS  
XX Claim 1: Page 59; 118bp; English.  
XX  
CC The present sequence encodes a protein from Streptococcus pneumoniae.  
CC The nucleic acid sequence encoding the Streptococcus pneumoniae protein  
CC can be useful in vaccines for inducing protective antibodies against  
CC Streptococcus pneumoniae, for treatment or prevention of infection e.g.  
CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid  
CC are used to detect Streptococcus infection (by usual hybridisation or  
CC amplification methods), also for isolating Streptococcus genes or their  
CC allelic variants. The protein can be used similarly for detecting specific  
CC antibodies in standard immunoassays, especially for diagnosing or  
CC monitoring infections. Antibodies which bind the protein are used to  
CC detect corresponding antigens, to purify the protein and for passive  
CC immunisation (optionally coupled to a toxin). Vaccines are administered,  
CC e.g. by injection, orally or through the skin, typically at 0.01-1000  
CC (especially 10-300) mu g/ml per dose.  
XX  
SQ Sequence 2389 BP; 830 A; 461 C; 486 G; 611 T; 1 other:  
Query Match 100.0%; Score 2388; DB 19; Length 2389;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|||||  
Db ATCAGCTAGAGATTGGCTGTGCAGAACGCTTCTATCTGGTGCAGGAATCTGTCAAA 660  
QY 601  
661 TTCGAACACCTATCGCGGCAAAAATAGCGATACACTTCAAGAACAACTGGGTACTTC 720  
Db 661  
721 TTCGAACACCTATCGCGGCAAAAATAGCGATACACTTCAAGAACAACTGGGTACTTC 720  
QY 721  
722 TGTAAACAATCCAGAACATACAAATCTAACACCAACCAACAGCACTAACAGTCA 780  
Db 721  
781 TGTAAACAATCCAGAACATACAAATCTAACACCAACCAACAGCACTAACAGTCA 780  
QY 781  
782 AGCAAGTCAAAGTATGACATTTAGTCTCTTGAAGACGCTCTACAAAGCTTTGAG 840  
Db 781  
841 AGCAAGTCAAAGTATGACATTTAGTCTCTTGAAGACGCTCTACAAAGCTTTGAG 840  
QY 841  
842 TCACACACATGTATGATCTGATGAGCTTGTCTTGTATCCAGACAAATACAAAGTGAAC 900  
Db 841  
901 TCACACACATGTATGATCTGATGAGCTTGTCTTGTATCCAGACAAATACAAAGTGAAC 900  
QY 901  
902 AGCTAGAGGTGTGTCAGTGCACACAGGAGATCATACCACTTCATCCCTTACTCTCAAT 960  
Db 901  
961 AGCTAGAGGTGTGTCAGTGCACACAGGAGATCATACCACTTCATCCCTTACTCTCAAT 960  
QY 961  
962 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1020  
Db 961  
1021 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1020  
QY 1021  
1022 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1080  
Db 1021  
1081 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1080  
QY 1081  
1082 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1140  
Db 1081  
1141 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1140  
QY 1141  
1142 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1200  
Db 1141  
1201 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1200  
QY 1201  
1202 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1260  
Db 1201  
1261 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1320  
QY 1261  
1262 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1320  
Db 1261  
1321 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1380  
QY 1321  
1381 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1440  
Db 1381  
1441 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1440  
QY 1441  
1442 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1500  
Db 1441  
1501 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1500  
QY 1501  
1502 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1560  
Db 1501  
1561 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1620  
QY 1561  
1621 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1680  
Db 1621  
1681 GTCTGATTTGGAAGAACGATGCTGTATTTATCCCTTGTCTTATCGTTCAACCATTTG 1740  
QY 1681

Db 1681 CTTTTCTGATAGAAAAAGTTGACGCTCAAGCTTATCTAAGAAAAAGTATCCACC 1740  
 QY 1741 TCCATCTCCAGACGAGATGTATAAGCAATCCAACTGAGATAGTCACAGCTATTTTA 1800  
 Db 1741 TCCATCTCCAGACGAGATGTATAAGCAATCCAACTGAGATAGTCACAGCTATTTTA 1800  
 QY 1801 CAATCGTGTGAAAGGGGAAAAAGCAATTCACGTCGTCGATTCGATATATGTTGAGCA 1860  
 Db 1801 CAATCGTGTGAAAGGGGAAAAAGCAATTCACGTCGTCGATTCGATATATGTTGAGCA 1860  
 QY 1861 TACAGTTGAGGTTAAAAACGGTATTTGATTTATTCCTCATTAAGATCATTAATAT 1920  
 Db 1861 TACAGTTGAGGTTAAAAACGGTATTTGATTTATTCCTCATTAAGATCATTAATAT 1920  
 QY 1921 TAAATTTGCTGGTTTGATGATACACATACAAAGCTCCAAATGGCTATACCTTTGAGAA 1980  
 Db 1921 TAAATTTGCTGGTTTGATGATACACATACAAAGCTCCAAATGGCTATACCTTTGAGAA 1980  
 QY 1981 TTTGTTTGCAGCATTTAAGTACTAGTAGAACAACCTGACGAACGTCACATTCCTAATGA 2040  
 Db 1981 TTTGTTTGCAGCATTTAAGTACTAGTAGAACAACCTGACGAACGTCACATTCCTAATGA 2040  
 QY 2041 TGGATGGGGGAATGCGACGTAGCATGTGTAGGCAAGAAAGACACAGTAAATCCAAA 2100  
 Db 2041 TGGATGGGGGAATGCGACGTAGCATGTGTAGGCAAGAAAGACACAGTAAATCCAAA 2100  
 QY 2101 TAAAGACTTCAAAAGCGGATGAAGAGCCAGTAGAGAAACACCTGCTGAGCAGAGTCCC 2160  
 Db 2101 TAAAGACTTCAAAAGCGGATGAAGAGCCAGTAGAGAAACACCTGCTGAGCAGAGTCCC 2160  
 QY 2161 TCAAGTAGAGACTGAAAAAGTAGAAGCCCAACTCAAAAGAGAGAAATTTTGTGGCAA 2220  
 Db 2161 TCAAGTAGAGACTGAAAAAGTAGAAGCCCAACTCAAAAGAGAGAAATTTTGTGGCAA 2220  
 QY 2221 AGTAAGGAGATTCAGTCTGGAAGCCATGCAACAGAAACTCTAGCTGTTAGGAATTA 2280  
 Db 2221 AGTAAGGAGATTCAGTCTGGAAGCCATGCAACAGAAACTCTAGCTGTTAGGAATTA 2280  
 QY 2281 TTTGACTCTTCAATATTATGATTAACAATAGTATCATATGCGCAAGAACGAAAAATTACTTGC 2340  
 Db 2281 TTTGACTCTTCAATATTATGATTAACAATAGTATCATATGCGCAAGAACGAAAAATTACTTGC 2340  
 QY 2341 GTTGTAAAAAGAGATTAATCTTCATCTGTAGTAAGAAAAATAAAC 2389  
 Db 2341 GTTGTAAAAAGAGATTAATCTTCATCTGTAGTAAGAAAAATAAAC 2389  
 RESULT 2  
 ABQ84819 ID ABQ84819 standard: DNA; 2389 BP.  
 AC ABQ84819;  
 XX 04-SEP-2002 (first entry)  
 DT S. pneumoniae SP036 nucleotide sequence SEQ ID NO:55.  
 DE Streptococcus pneumoniae: epitope; vaccine; antigenic protein;  
 KM antibacterial; Streptococcal infection; detection; gene; ds.  
 XX Streptococcus pneumoniae.  
 OS US2002061545-A1.  
 PN 23-MAY-2002.  
 PD 22-JAN-2001; 2001US-0765272.  
 PF 30-OCT-1997; 97US-0961083.  
 PR (CHOI/) CHOI G H.  
 PA (KUNSH/) KUNSH C A.  
 PA (BARA/) BARASH S C.

PA (DILL/) DILLON P J.  
 PA (DOUG/) DOUGHERTY B.  
 PA (FANN/) FANNON M R.  
 PA (ROSE/) ROSEN C A.  
 PI Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR;  
 PI Rosen CA;  
 DR WPI: 2002-479261/51.  
 DR P-PSDB; ABP54584.  
 XX New Streptococcus pneumoniae antigens, useful for detecting  
 PT Streptococcus and for preventing or attenuating disease caused by  
 PT Streptococcus infection -  
 XX  
 PS Claim 1; Page 27; 70pp; English.  
 CC ABQ84792 to ABQ84904 represents nucleic acids which encode the  
 CC Streptococcus pneumoniae antigens given in ABP54557 to ABP54669.  
 CC The S. pneumoniae antigens have antibacterial activity and can be  
 CC used in vaccines. The S. pneumoniae antigens can also be used to  
 CC prevent or attenuate a Streptococcal infection in an animal. The  
 CC polynucleotides encoding the S. pneumoniae antigens can be used to  
 CC detect Streptococcus nucleic acids. ABQ84905 to ABQ85130 represent  
 CC primers used in the cloning of S. pneumoniae ORFs (open reading frames)  
 CC which are used in an example from the present invention.  
 XX  
 SQ Sequence 2389 BP; 830 A; 461 C; 486 G; 611 T; 1 other;  
 Query Match 100.0%; Score 2388; DB 24; Length 2389;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TTTCTTACGATTTGGGATGTATCAAGCTAGACGCTTAAGAAAAATATGTTCTTA 60  
 Db 1 TTTCTTACGATTTGGGATGTATCAAGCTAGACGCTTAAGAAAAATATGTTCTTA 60  
 QY 61 TATAGATGAAAAACAACGACGCAAAAAACGAGAAATTTGACCTCATGAGTTAGCAA 120  
 Db 61 TATAGATGAAAAACAACGACGCAAAAAACGAGAAATTTGACCTCATGAGTTAGCAA 120  
 QY 121 GCGTGAAGGATCAATGCTGAGCAAAATCGATCAAGATTAACAGACCAGCTATGTCAC 180  
 Db 121 GCGTGAAGGATCAATGCTGAGCAAAATCGATCAAGATTAACAGACCAGCTATGTCAC 180  
 QY 181 TTTACATGCGGACACATATATTTACATGGAAGTCTTATATACGCTATCATCAG 240  
 Db 181 TTTACATGCGGACACATATATTTACATGGAAGTCTTATATACGCTATCATCAG 240  
 QY 241 TGAAGATTTCTCATGAAAGATCCAACTTAAGCTAAAGATGAGATATTTGTTAATGA 300  
 Db 241 TGAAGATTTCTCATGAAAGATCCAACTTAAGCTAAAGATGAGATATTTGTTAATGA 300  
 QY 301 GGTCAAGGATGATATGTTATCAAGGTAGATGAAAAATATATCTTAAAGATGC 360  
 Db 301 GGTCAAGGATGATATGTTATCAAGGTAGATGAAAAATATATCTTAAAGATGC 360  
 QY 361 TGCCACAGCGGATRACCTCGTACAAAAGAGAAATATATGACAAAACAGAGCATAG 420  
 Db 361 TGCCACAGCGGATRACCTCGTACAAAAGAGAAATATATGACAAAACAGAGCATAG 420  
 QY 421 TCAACATCGTGAAGGTGGAATCCCAAGAAAGATGGTGGCTTGACAGTTGCA 480  
 Db 421 TCAACATCGTGAAGGTGGAATCCCAAGAAAGATGGTGGCTTGACAGTTGCA 480  
 QY 481 AGGACGCTATACACAGATGATGTTATATCTTAATGCTTGTATATACATAGAGATAC 540  
 Db 481 AGGACGCTATACACAGATGATGTTATATCTTAATGCTTGTATATACATAGAGATAC 540  
 QY 541 TGGTGAATGCTTATATGCTTCTCATGAGATCATTTACCATTAATCTTAAGATGATT 600  
 Db 541 TGGTGAATGCTTATATGCTTCTCATGAGATCATTTACCATTAATCTTAAGATGATT 600



QY 601 ATCAGCTAGCGAGTGGCTGCTGCTGAGAGCCCTTCTATCTGCTGAGGAATCTGTCAA 660  
 Db 601 ATCAGCTAGCGAGTGGCTGCTGCTGAGAGCCCTTCTATCTGCTGAGGAATCTGTCAA 660  
 QY 661 TTCAAGACCTATCGCCGACAAATATAGCGATACACTTCAAGAACCACTGGGTACTTTC 720  
 Db 661 TTCAAGACCTATCGCCGACAAATATAGCGATACACTTCAAGAACCACTGGGTACTTTC 720  
 QY 721 TGTAAACAATCCAGAACTCAATTAACACAGCAACAGCAACCACTAACAGTCA 780  
 Db 721 TGTAAACAATCCAGAACTCAATTAACACAGCAACAGCAACCACTAACAGTCA 780  
 QY 781 AGCAAGTCAAAAGTATGACATTTAGTCTCTTGAACACAGCTCTAGCAACTGGCTTTGAG 840  
 Db 781 AGCAAGTCAAAAGTATGACATTTAGTCTCTTGAACACAGCTCTAGCAACTGGCTTTGAG 840  
 QY 841 TCACAGCAATGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 Db 841 TCACAGCAATGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 QY 901 AGCTAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 Db 901 AGCTAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 QY 961 GTCTGATTTGGAAGAAAGAAATGCTGATTTATCCCTTCTGTTATCTGTTCAACCAATTG 1020  
 Db 961 GTCTGATTTGGAAGAAAGAAATGCTGATTTATCCCTTCTGTTATCTGTTCAACCAATTG 1020  
 QY 1021 GGTACCAAGTCAAGGCGCAACCAACCACTGCAACCAACCACTGCAACCACTGCAACCA 1080  
 Db 1021 GGTACCAAGTCAAGGCGCAACCAACCACTGCAACCAACCACTGCAACCACTGCAACCA 1080  
 QY 1081 CCGGCAACCTGCAACCAATCTTAAATAGACTCAATTTCTTGTGTTAGTCAAGTGTG 1140  
 Db 1081 CCGGCAACCTGCAACCAATCTTAAATAGACTCAATTTCTTGTGTTAGTCAAGTGTG 1140  
 QY 1141 ACAGAAAGTTGGGGAAGATATGTTTCAAGAAAGGCACTGCTGCTGCTGCTGCTGCTGCT 1200  
 Db 1141 ACAGAAAGTTGGGGAAGATATGTTTCAAGAAAGGCACTGCTGCTGCTGCTGCTGCTGCT 1200  
 QY 1201 GAAAGATTTACATCTGAAACCTGTTAAATCTGAAAGCAATTTCAAAACCAAGAGAG 1260  
 Db 1201 GAAAGATTTACATCTGAAACCTGTTAAATCTGAAAGCAATTTCAAAACCAAGAGAG 1260  
 QY 1261 TGTTCACACACTTTAACTGCTTAAAGAAAGAAATGTTGCTGCTGCTGCTGCTGCTGCT 1320  
 Db 1261 TGTTCACACACTTTAACTGCTTAAAGAAAGAAATGTTGCTGCTGCTGCTGCTGCTGCT 1320  
 QY 1321 TGTAAAGCATTAATCTGTTAACTGAGGCTCAATAAGCCTTTGTTGAAATTAAGGCTG 1380  
 Db 1321 TGTAAAGCATTAATCTGTTAACTGAGGCTCAATAAGCCTTTGTTGAAATTAAGGCTG 1380  
 QY 1381 TAATTTGATTTCCAGACCTTTAGCAAAATTAATAGAAAGCTTTGTTGAAATTAAGGCTG 1440  
 Db 1381 TAATTTGATTTCCAGACCTTTAGCAAAATTAATAGAAAGCTTTGTTGAAATTAAGGCTG 1440  
 QY 1441 TAAAGAAATTTGATGATGATTTATTTGCAATTTCTAGACCAATTAACCAATCAAGGCG 1500  
 Db 1441 TAAAGAAATTTGATGATGATTTATTTGCAATTTCTAGACCAATTAACCAATCAAGGCG 1500  
 QY 1501 ACTTGGCAACCAATTTCTCAATTTGATGATGATGATGATGATGATGATGATGATGATG 1560  
 Db 1501 ACTTGGCAACCAATTTCTCAATTTGATGATGATGATGATGATGATGATGATGATGATG 1560  
 QY 1561 AGCTGATTAAGTATACACGTCAGATGTTTATTTGATGATGATGATGATGATGATGATGATG 1620  
 Db 1561 AGCTGATTAAGTATACACGTCAGATGTTTATTTGATGATGATGATGATGATGATGATGATG 1620  
 QY 1621 TGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680  
 Db 1621 TGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680  
 QY 1681 CTTTCTGATTAAGGAAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740

Db 1681 CTTTCTGATTAAGGAAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
 QY 1741 TCCATCTCCAGACGATGATTTAAAGCAATTCACCTGAGATGATGATGATGATGATGATGAT 1800  
 Db 1741 TCCATCTCCAGACGATGATTTAAAGCAATTCACCTGAGATGATGATGATGATGATGATGAT 1800  
 QY 1801 CAATGCTGTGAAAGGGGAAAGAAACGATTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860  
 Db 1801 CAATGCTGTGAAAGGGGAAAGAAACGATTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860  
 QY 1861 TACAGTTGAGGTTAAAGGATTTGATTTATTTCTGATGATGATGATGATGATGATGATGAT 1920  
 Db 1861 TACAGTTGAGGTTAAAGGATTTGATTTATTTCTGATGATGATGATGATGATGATGATGAT 1920  
 QY 1921 TAAATTTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980  
 Db 1921 TAAATTTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980  
 QY 1981 TTTGTTGGGACGATTTAAGTACTAGCTAGACACCTGAGCAAGCTGCAATTTCTAATGA 2040  
 Db 1981 TTTGTTGGGACGATTTAAGTACTAGCTAGACACCTGAGCAAGCTGCAATTTCTAATGA 2040  
 QY 2041 TGGATGGGCAATGCGAGTGAAGCATGTTAGGCAAGAAAGACACAGTGAAGATCCAA 2100  
 Db 2041 TGGATGGGCAATGCGAGTGAAGCATGTTAGGCAAGAAAGACACAGTGAAGATCCAA 2100  
 QY 2101 TAAAGCTTCAAGGCGATGAAGAGGCGATGAGGAAACACCTGCTGAGCCAGAAATGCC 2160  
 Db 2101 TAAAGCTTCAAGGCGATGAAGAGGCGATGAGGAAACACCTGCTGAGCCAGAAATGCC 2160  
 QY 2161 TCAAGTAGAGACTGAAAGAAAGTGAAGCCCACTCAAGAAAGGAAAGTTGCTTGCAGAA 2220  
 Db 2161 TCAAGTAGAGACTGAAAGAAAGTGAAGCCCACTCAAGAAAGGAAAGTTGCTTGCAGAA 2220  
 QY 2221 AGTAAAGGATTTAGTCTGAAAGCCCAATGCAACAGAAACTTATGCTGCTTACGAATTA 2280  
 Db 2221 AGTAAAGGATTTAGTCTGAAAGCCCAATGCAACAGAAACTTATGCTGCTTACGAATTA 2280  
 QY 2281 TTTGACTCTTCAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340  
 Db 2281 TTTGACTCTTCAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340  
 QY 2341 GTTGTAAAGGAAGTAACTCTTCACTGTTAAGTAAAGGAAAGAAATTAAC 2389  
 Db 2341 GTTGTAAAGGAAGTAACTCTTCACTGTTAAGTAAAGGAAAGAAATTAAC 2389  
 RESULT 3  
 AAA47604  
 ID AAA47604 standard; DNA; 2451 BP.  
 XX  
 AC AAA47604;  
 XX  
 DT 20-OCT-2000 (first entry)  
 XX  
 DE Recombinant variant of Sp36 gene (Sp36A) of *S. pneumoniae*.  
 XX  
 KW Streptococcus pneumoniae; infection; vaccine; collid coil region;  
 KW histidine triad residue; Sp36; antibody; otitis media;  
 KW nasopharyngeal infection; bronchial infection; bronchitis; sepsis;  
 KW meningitis; lobar pneumonia; ds.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..2451  
 FT CDS /\*tag= a  
 FT /product= Sp36a polypeptide  
 PN MO200037105-A2.  
 XX  
 PD 29-JUN-2000.



```
OY 1621 TGAAGGATGATATGTAAACGGCTCATATATGGCCATAGTCACTGGATTGGAAAAGATAG 1680
    |||||||
DB 1680 TGAAGGATGATATGTAAACGGCTCATATATGGCCATAGTCACTGGATTGGAAAAGATAG 1739
OY 1681 CCTTCTGATAGAAAAAGTTGCAGCTCAAGCCTATACTAAAGAAAAAGGTATCCACC 1740
    |||||||
DB 1740 CCTTCTGATAGAAAAAGTTGCAGCTCAAGCCTATACTAAAGAAAAAGGTATCCACC 1799
OY 1741 TCCATCTCCAGAGCAGATGTTAAAGCAATCCAACTGAGATAGTGCACACTATTTA 1800
    |||||||
DB 1800 TCCATCTCCAGAGCAGATGTTAAAGCAATCCAACTGAGATAGTGCACACTATTTA 1859
OY 1801 CAATCTGTGAAAGGGGAAAAACGAATTCACATCGTTGCACTTCATATATGTTGAGCA 1860
    |||||||
DB 1860 CAATCTGTGAAAGGGGAAAAACGAATTCACATCGTTGCACTTCATATATGTTGAGCA 1919
OY 1861 TACAGTTGAGTTAAAAACGGTAAATTGATTTATTCCTCATAGGATCATTTACATAATAT 1920
    |||||||
DB 1920 TACAGTTGAGTTAAAAACGGTAAATTGATTTATTCCTCATAGGATCATTTACATAATAT 1979
OY 1921 TAAATTGCTGGTTGATGATACACATACAAAGCTCCAAATGGGTATCCCTGGAGA 1980
    |||||||
DB 1980 TAAATTGCTGGTTGATGATACACATACAAAGCTCCAAATGGGTATTCCTGGAGA 2039
OY 1981 TTTGTTGCGAGATTAAAGTACTACGTAGAACACCCGTGAGAACGTCCACATTTCTAATGA 2040
    |||||||
DB 2040 TTTGTTGCGAGATTAAAGTACTACGTAGAACACCCGTGAGAACGTCCACATTTCTAATGA 2099
OY 2041 TGGATGGGGAATGCGCAGTAGAGCATGTGTTAGCAAGAAAGACCACAGTAAATCCAAA 2100
    |||||||
DB 2100 TGGATGGGGAATGCGCAGTAGAGCATGTGTTAGCAAGAAAGACCACAGTAAATCCAAA 2159
OY 2101 TAAGACTTCAAAACGGATGAAGAGCCAGTAGAGAAAACCTGCTGAGCCAGAGTCCC 2160
    |||||||
DB 2160 TAAGACTTCAAAACGGATGAAGAGCCAGTAGAGAAAACCTGCTGAGCCAGAGTCCC 2219
OY 2161 TCAAGTAGACACTGAAAAAGTAGAAGCCCACTCAAAAGAAAGAGATTGCTTGCAGAA 2220
    |||||||
DB 2220 TCAAGTAGACACTGAAAAAGTAGAAGCCCACTCAAAAGAAAGAGATTGCTTGCAGAA 2279
OY 2221 AGTAAGGATTTAGTCTGAAAGCCAAATGCAAGAAACCTGCTGCTTAGGAATAA 2280
    |||||||
DB 2280 AGTAAGGATTTAGTCTGAAAGCCAAATGCAAGAAACCTGCTGCTTAGGAATAA 2339
OY 2281 TTTGACTCTTCAATATATGATTAACAATAGTATCATGAGCAGAGAAAAATTTACTTGC 2340
    |||||||
DB 2340 TTTGACTCTTCAATATATGATTAACAATAGTATCATGAGCAGAGAAAAATTTACTTGC 2399
OY 2341 GTTGTAAAAAGAAATATCTTTCATCTGTAAAGTAAAGAAAAATTAAC 2389
    |||||||
DB 2400 GTTGTAAAAAGAAATATCTTTCATCTGTAAAGTAAAGAAAAATTAAC 2448

RESULT 4
AAV52227
ID AAV52227 standard; DNA; 8195 BP.
XX
AC AAV52227;
XX
DE 23-OCT-1998 (first entry)
XX
KW Streptococcus pneumoniae, genome fragment SEQ ID NO:94.
XX
KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
computer readable medium; vaccine; pharmaceutical composition; ds.
XX
OS Streptococcus pneumoniae.
XX
PN MO9818931-A2.
XX
PD 07-MAY-1998.
XX
```

```
PF 30-OCT-1997; 97WO-US19588.
XX
XX 31-OCT-1996; 96US-0028960.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
PI Kunsch CA, Rosen CA;
XX
XX MPI; 1998-272225/24.
XX
XX Computer-readable medium with recorded Streptococcus pneumoniae
PT polynucleotide sequences - useful in diagnostic kits and assays, and
PT pharmaceutical compositions and vaccines for Streptococcus
XX pneumoniae
XX
PS Claim 1: Page 727-732; 1409pp; English.
XX
XX The present invention describes a computer readable medium which has
CC the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)
CC recorded on it, or a representative fragment or a sequence at least 95%
CC identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in
CC SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from
CC Streptococcus pneumoniae. The present invention also describes an
CC isolated nucleic acid molecule encoding a homologue of any of the
CC fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the
CC nucleic acid molecule is produced by a process comprising: (a) screening
CC a genomic DNA library using as a probe a target sequence defined by any
CC of the sequences in SEQ ID NO:1 to 391, identifying members of the
CC library which contain sequences that hybridise to the target sequence and
CC isolating the nucleic acid molecules from the members; or (b) isolating
CC mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
CC molecules whose nucleotide sequence is homologous to amplification
CC primers derived from the fragment of the S. pneumoniae genome to prime
CC the amplification and isolating the amplified sequences. The computer
CC readable medium can be used in a computer-based system for identifying
CC fragments of the S. pneumoniae genome of commercial importance, or
CC expression modulating fragments of the S. pneumoniae genome. Products
CC from the present invention can be used in diagnosis kits and assays, and
CC pharmaceutical compositions and vaccines for S. pneumoniae.
XX
XX Sequence 8195 BP; 2688 A; 1622 C; 1777 G; 2105 T; 3 other:
XX
XX Query Match 100.0%; Score 2388; DB 19; Length 8195;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 2388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTTCTTACGAGTTGGGACTGTATCAAGCTAGAACGGTTAAGAAAAATATCGTTTCCTA 60
    |||||||
DB 3053 TTCTTACGAGTTGGGACTGTATCAAGCTAGAACGGTTAAGAAAAATATCGTTTCCTA 3112
OY 61 TATAGATGAAAAACAAGCGACGCAAAAAACGGAGATTTGACCTCCGATGAGTTTGCATA 120
    |||||||
DB 3113 TATAGATGAAAAACAAGCGACGCAAAAAACGGAGATTTGACCTCCGATGAGTTTGCATA 3172
OY 121 GCGTGAAGGAATCAATGCTGAGCAAAATCGTCATCAAGATACAGACCAAGCGTATGTCAC 180
    |||||||
DB 3173 GCGTGAAGGAATCAATGCTGAGCAAAATCGTCATCAAGATACAGACCAAGCGTATGTCAC 3232
OY 181 TTCACATGGCGACACTATCATTTATTACAATGGTAAAGTTCTTATGACGCTATCATCAG 240
    |||||||
DB 3233 TTCACATGGCGACACTATCATTTATTACAATGGTAAAGTTCTTATGACGCTATCATCAG 3292
OY 241 TGAAGATATTCATGAAAGATCCAAACATTAAGCTAAAGAAAGTGAATGTATTAATA 300
    |||||||
DB 3293 TGAAGATATTCATGAAAGATCCAAACATTAAGCTAAAGAAAGTGAATGTATTAATA 3352
OY 301 GGTCAAGGCTGATATGTATCAAGGTAGATGAAATATCTATGTTTACCTTAAGGATGC 360
    |||||||
DB 3353 GGTCAAGGCTGATATGTATCAAGGTAGATGAAATATCTATGTTTACCTTAAGGATGC 3412
OY 361 TGGCCACGGCGATTAACGTCCGTACAAAGAGAAATCAATGACAAAAAAGAGCATAG 420
    |||||||
```

D	3413	TGCCACGCCGATATACCTCCGTCACAAAAGGAAATCATCGACAAAAACAAGACATAG	3472
O	421	TCACATCTGTGAAGTGGAACTCCAGAAAAGATGGTGGCTGTGGCACTTGCGCA	480
D	3473	TCACATCTGTGAAGTGGAACTCCAGAAAAGATGGTGGCTGTGGCACTTGCGCA	3532
O	481	AGGACGCTATCTACACATATGCTTATCTTTAATGCTTGATATCATAGAGATAC	540
D	3533	AGGACGCTATCTACACATATGCTTATCTTTAATGCTTGATATCATAGAGATAC	3592
O	541	TGGATGAGCTTATATGCTTCCATGAGAGATCTTACCTTCAATCTCCAGAAAGATTT	600
D	3593	TGGATGAGCTTATATGCTTCCATGAGAGATCTTACCTTCAATCTCCAGAAAGATTT	3652
O	601	ATCAGCTAGCAGATTGGCTGCTGACAGAGCCTTCTATGTGGTCAGAGAAATCTGTCAA	660
D	3653	ATCAGCTAGCAGATTGGCTGCTGACAGAGCCTTCTATGTGGTCAGAGAAATCTGTCAA	3712
O	661	TTCAAGAACCTATATGGCGACAAAATAGCGATTAACCTTAAGACAAACTGGGTACCTTC	720
D	3713	TTCAAGAACCTATATGGCGACAAAATAGCGATTAACCTTAAGACAAACTGGGTACCTTC	3772
O	721	TGTAAAGCATTCAGAGACTACAAATACTAACACACAAACAGACACACACTAACAGTCA	780
D	3773	TGTAAAGCATTCAGAGACTACAAATACTAACACACAAACAGACACACACTAACAGTCA	3832
O	781	AGCAAGTCAAAGATATGATACATTTAGTGTCTTGAAACAGCTCTACAAATCTGCTTTGAG	840
D	3833	AGCAAGTCAAAGATATGATACATTTAGTGTCTTGAAACAGCTCTACAAATCTGCTTTGAG	3892
O	841	TCACAGCATCTAGATCTGATGCGCTTGTCTTTGATCCAGACAAATACACAGTGCAC	900
D	3893	TCACAGCATCTAGATCTGATGCGCTTGTCTTTGATCCAGACAAATACACAGTGCAC	3952
O	901	AGCTAGAGGTTTGCATGTGCGCACACGAGATCTTACCACTTCAATCCCTACTCTCAAT	960
D	3953	AGCTAGAGGTTTGCATGTGCGCACACGAGATCTTACCACTTCAATCCCTACTCTCAAT	4012
O	961	GCTCGAATTGGAAAGAGAAATGCGCTGATATTTCCCTCTGTAATCGTTAAACCATTG	1020
D	4013	GCTCGAATTGGAAAGAGAAATGCGCTGATATTTCCCTCTGTAATCGTTAAACCATTG	4072
O	1021	GGTACACAGATTCAAGGCCAGACACACACAGTCCGAGACCTGAGCCTAGTCCAGG	1080
D	4073	GGTACACAGATTCAAGGCCAGACACACAGTCCGAGACCTGAGCCTAGTCCAGG	4132
O	1081	CCCGCAACCTGCAACCAATCTTAAATTAAGCTCAATTTCTTTGGTTAGTCACTGCTGT	1140
D	4133	CCCGCAACCTGCAACCAATCTTAAATTAAGCTCAATTTCTTTGGTTAGTCACTGCTGT	4192
O	1141	ACGAAAAGTTGGGAGAGATATGATTTGAGAAAAGGCACTCTCGTATGCTTTGCG	1200
D	4193	ACGAAAAGTTGGGAGAGATATGATTTGAGAAAAGGCACTCTCGTATGCTTTGCG	4252
O	1201	GAAGATTTACCATCTGAAACTGTAAAAATCTTGAAGCAAGTTATCAAAACAAGAGAG	1260
D	4253	GAAGATTTACCATCTGAAACTGTAAAAATCTTGAAGCAAGTTATCAAAACAAGAGAG	4312
O	1261	TGTTTCACACCTTTACCTCTATAAAAAGAAAATGTGCTCTCTGTCACAGAATTTTA	1320
D	4313	TGTTTCACACCTTTACCTCTATAAAAAGAAAATGTGCTCTCTGTCACAGAATTTTA	4372
O	1321	TGATTAAGCATATATCTGTTAACTAGAGCTCATTAAGCCTTGTTGNAATAAAGGGTCG	1380
D	4373	TGATTAAGCATATATCTGTTAACTAGAGCTCATTAAGCCTTGTTGNAATAAAGGGTCG	4432
O	1381	TAAATTGATTTCCAAAGCCTTAGACAAATTTAGAACCTTGAATGATGCAGCTAA	1440
D	4433	TAAATTGATTTCCAAAGCCTTAGACAAATTTAGAACCTTGAATGATGCAGCTAA	4492
O	1441	TAAAGAAAATTTGGTATATATTTATTGGCATTTCTAGACCAATTTACCATCCAGAGCG	1500
D	4493	TAAAGAAAATTTGGTATATATTTATTGGCATTTCTAGACCAATTTACCATCCAGAGCG	4552

QY	1501	ACTGTCGAAACCAAAATTCCTCAAAATGTAGATATCTGAAAGACAGAAATTCGATATGCTCAATT	1560
Db	4553	ACTTGGCAAAACCAAAATTCCTCAAAATGTAGATATCTGAAAGACAGAAATTCGATATGCTCAATT	4612
QY	1561	AGCTATATAGATATCAACGTCAGATGGTTTCATTTTGTATGAACATGATATATACAGA	1620
Db	4613	AGCTATATAGATATCAACGTCAGATGGTTTCATTTTGTATGAACATGATATATACAGA	4672
QY	1621	TGAAGAGATGCATATGTATTAACGGCTCATATATGGGCCATAGTCACTGATTTGAAAAATAG	1680
Db	4673	TGAAGAGATGCATATGTATTAACGGCTCATATATGGGCCATAGTCACTGATTTGAAAAATAG	4732
QY	1681	CCTTTCGTATATAGAAAAAAGTTGACGCTCAAGCCTTACTATAAGAAAAAGTATCTTACC	1740
Db	4733	CCTTTCGTATATAGAAAAAAGTTGACGCTCAAGCCTTACTATAAGAAAAAGTATCTTACC	4792
QY	1741	TCCATCTCCAAACCCAGATGTTAAAGCAAAATCCAACTGAGATATGTCAGCAGTATTTA	1800
Db	4793	TCCATCTCCAAACCCAGATGTTAAAGCAAAATCCAACTGAGATATGTCAGCAGTATTTA	4852
QY	1801	CAATCGTGTGAAGGGGAAAAAGCAATTCGCTGTTGCACTTCATATATGTTGAGCA	1860
Db	4853	CAATCGTGTGAAGGGGAAAAAGCAATTCGCTGTTGCACTTCATATATGTTGAGCA	4912
QY	1861	TACAGTTAGGTTTAAAAAGGTTAATTTGATTTCTCTCAATAGAAATTCATTAACATATAT	1920
Db	4913	TACAGTTAGGTTTAAAAAGGTTAATTTGATTTCTCTCAATAGAAATTCATTAACATATAT	4972
QY	1921	TAAATTTGCTGTTGTTGATATATCAACATCAAAAGCTCCAAATGGCTATACCTTGGAAGA	1980
Db	4973	TAAATTTGCTGTTGTTGATATATCAACATCAAAAGCTCCAAATGGCTATACCTTGGAAGA	5032
QY	1981	TTTGTGTTCCGACGTTTAAGTACTACGTAGAACACCTGACGAAGCTCCACATTTCTAATGA	2040
Db	5033	TTTGTGTTCCGACGTTTAAGTACTACGTAGAACACCTGACGAAGCTCCACATTTCTAATGA	5092
QY	2041	TGGATGGGGCAATGCCAGTGGGACATGTTTAGGCAAGAAAGCAACAGTGAAGATCCAAA	2100
Db	5093	TGGATGGGGCAATGCCAGTGGGACATGTTTAGGCAAGAAAGCAACAGTGAAGATCCAAA	5152
QY	2101	TAAGAACTTCAAAGCGGATGAAGAGCCAGTAGAGGAACACCTGCTGAGCCAGAAAGTCCC	2160
Db	5153	TAAGAACTTCAAAGCGGATGAAGAGCCAGTAGAGGAACACCTGCTGAGCCAGAAAGTCCC	5212
QY	2161	TCAGTATAGACGCTAAAAAGTAGAGCCCACTCAAAAGCAGAGAAATTTGCTTGGCAA	2220
Db	5213	TCAGTATAGACGCTAAAAAGTAGAGCCCACTCAAAAGCAGAGAAATTTGCTTGGCAA	5272
QY	2221	AGTAAAGGATTCGTGCTGGAAGCCAAATGCAACAGAAACTATAGCTGGTTTACGAAATTA	2280
Db	5273	AGTAAAGGATTCGTGCTGGAAGCCAAATGCAACAGAAACTATAGCTGGTTTACGAAATTA	5332
QY	2281	TTTGACCTCTTCAAAATTTATGATTAACAATAGTATCATGCGCAGAGACGAAAAATTTACTTGC	2340
Db	5333	TTTGACCTCTTCAAAATTTATGATTAACAATAGTATCATGCGCAGAGACGAAAAATTTACTTGC	5392
QY	2341	GTTGTTAAAGGAAGTATATCTTCATCTGTAACTAAGAAAAAATTAAC	2389
Db	5393	GTTGTTAAAGGAAGTATATCTTCATCTGTAACTAAGAAAAAATTAAC	5441
RESULT 5			
AAA65731			
ID	AAA65731 standard; DNA: 2523 BP.		
XX	AAA65731;		
XX	21-NOV-2000 (first entry)		
XX	Streptococcus pneumoniae BVH-11 gene SEQ ID NO:3.		
XX	Streptococcus pneumoniae; BVH-3; BVH-11; BVH-28; antigen; vaccine:		
KX	Streptococcus pneumoniae; BVH-3; BVH-11; BVH-28; antigen; vaccine:		



```

Db      1620 AGTAGCCAACTTGGCAGGCAAGTACACACAGAGAGGTTATATCTTTATCCTCGTGA 1679
OY      1609 TATTAATGAGTGAAGAGATGATATGTAAAGCCCTCATATGGCGATGCTACGTGAT 1668
Db      1680 TATTAACCGATGATGAGGGGATGCTATGTAACTCCATATATGACCATATGCGATGAT 1739
OY      1669 TGGAAAAGATAGCTCTTCTGTATAGAGAAAAGTTGACACTCAAGCCCTATCTAAAGAAA 1728
Db      1740 TAAAAAAGATAGTTTGTCTGTAAGCTGAGAGAGCGGACGCCAGGCTTATGCTAAAGAA 1799
OY      1729 AGGTATCTCTACCTCCATCTCCAGAGCAGATGTTAAAGCAATCCAACTGAGATAGTGC 1788
Db      1800 AGGTTTGACCCCTCCCTTCGACAGCCATCAGGATTCAGAAATATCTGAGCAAAAAGGAGC 1859
OY      1789 AGGAGCTATTTTACAAATGATGTAAGGGAAGAAAACGAATTCCTGCTGGACTCCATA 1848
Db      1860 AGAAGCTATCTTACACCGCTGAAAGCAGCTAAGAGAGTCCACTTATGATGATGCTTA 1919
OY      1849 TATGTTGACATACAGTGTAGGTTAAAAACGGTATTTGATTTCTCATTAAGATCA 1908
Db      1920 CAATCTTCAATATACGTGTAGAGTCAAAAACGGTATTTATCATACCTCATATATGACCA 1979
OY      1909 TTACCATTAATTAATTTGCTGTTGTTGATGATCACACATPACAAAGCTCCAAATGGCTA 1968
Db      1980 TTACCATTAACATCAAAATTTGATGTTGTTGACGAAGGCTTTATGAGGACACTAAGGGGTA 2039
OY      1969 TACCTTGAAGATTTGTTTGGACGATTAAGTACTAGTAAACACCCCTACAGAACTCC 2028
Db      2040 TACTCTTGAAGATCTTTTGGGACTGTCAAGTACTATGTGACATCAACGAAGACCTCC 2099
OY      2029 ACATTTCAATGATGATGGGCAATGCCAGTGAAGATGTTAGGCAAGAAGACCCAG 2088
Db      2100 GCATTCAGATTAATGGTTTGTAAACGCTAGCGACCATGTTCAAGAAACAAATAATGCTCA 2159
OY      2089 TGAAGATCCAAATPAGAACTTCAAAAGCGATGAGAG----- 2125
Db      2160 AGCTGATACCAATCAAAACGGAACCAACGAGAGAAACTCAGACAGAAAACCTGA 2219
OY      2126 ----- 2125
Db      2220 GGAAGAAACCCCTCGAGAGAGAAACCAAAAGGAGAAACAGAGTCTCCAAAACCAAC 2279
OY      2126 -----CCAGTAGAGAGAAACACCTGCTGAGCAGAAAGTCCCTCAAGTAGAGTGA 2178
Db      2280 AGAGGAACCGAAGAAAGAAATCCACAGAGAAATCAGAAAGAACTCAGTCCAGACTGAAA 2339
OY      2179 AGTAGAAGCCCAACTCAAGAGACGAAAGTTTGGCTTGCAGAAATGACGATTCCTAGCT 2238
Db      2340 GGTGGAAGAAAACCTGAGAGAGCTGAAAGATTACTTGAAAAATCCAGATCCAAATTA 2399
OY      2239 GAAGCAATGCAACAGAACTAGCTGTTACGAAATTAATTTGACTCTTCAATTA 2258
Db      2400 CAAGTCAATAGCCAAAGAGACTCTCAGAGATTAATAATTTACTATTTGGACCCA 2459
OY      2299 GGATTAACATAGTATCATGCGAGAGCAGAAAAATTTACTGCTGTTTAAAGGAAGTAA 2358
Db      2460 GGACAAACAATACATATATGCGAAGCTGAAAAACTATTTGGCTTTATTAAGGAGAGTAA 2519

```

RESULT 6  
ID AAA65736 standard: DNA: 2647 BP.  
XX  
AC AAA65736:  
XX

XX 21-NOV-2000 (first entry)

DE Streptococcus pneumoniae BVH-11 gene SEQ ID NO:12.

XX Streptococcus pneumoniae, BVH-3; BVH-11; BVH-28; antigen; vaccine;  
KW Proprietary; therapy; infection; diagnosis; meningitis; bacteraemia;  
KW Otitis media; pneumonia; immunisation; bactericidal; ds.

```

XX OS Streptococcus pneumoniae.
XX
XX MO200039299-A2.
XX
XX PD 06-JUL-2000.
XX
XX PF 20-DEC-1999; 99MO-CA01218.
XX
XX PR 23-DEC-1998; 98US-0113800.
XX
XX PA (BIOC-) BIOCHEM PHARMA INC.
XX
XX PI Hamel J, Brodeur BR, Pineau I, Martin D, Rioux C, Charland N,
XX WPI; 2000-452397/39.
XX
XX DR Streptococcal antigens useful for vaccinating against e.g. meningitis,
XX PT otitis media, bacteraemia and/or pneumonia -
XX PS Example 6; Fig 15; 106pp; English.
XX
XX CC The present invention describes nucleic acids (I) encoding protein
XX CC antigens (II) from Streptococcus pneumoniae. The protein antigens
XX CC have bactericidal activity. The nucleic acids, encoding the protein
XX CC antigens, may be used for the recombinant production of the proteins
XX CC they encode. The protein antigens may then be used as vaccines for the
XX CC prevention and treatment of Streptococcal infections in mammals
XX CC (especially humans) which result in, e.g. meningitis, otitis media,
XX CC bacteraemia and/or pneumonia. The present sequence encodes the
XX CC S. pneumoniae BVH-11 protein antigen.
XX
XX SQ Sequence 2647 BP; 934 A; 538 C; 556 G; 619 T; 0 other;
XX
XX Query Match 57.5%; Score 1374.2; DB 21; Length 2647;
XX Best Local Similarity 73.9%; Pred. No. 0;
XX Matches 1819; Conservative 0; Mismatches 539; Indels 102; Gaps 2;
OY 1 TTCTTAGAGTTGGGACTGTATCAAGCTAGACGGTTAAGGAAATTAATGCTTCTCTA 60
Db 104 TGTATAGAACTAGTTTGATCAAGCTCAAACTGTAAGAAATAATGCTGTTCTTA 163
OY 61 TATAGATGAAAACCAACGACGCAAAAACGAGAAATTTGACTCTGATAGTTAGCAA 120
Db 164 TATAGATGAAAACCAACGACGCAAAAACGAGAAATTTGACTCTGATAGTTAGCAA 223
OY 121 GCGTGAAGAAATCAATCTGAGCAAAATGCTCATCAAGATTAACAGACCAAGCTATGTCAC 180
Db 224 GCGTGAAGAAATCAACGCCGCAAAATGCTCATCAAGATTAACAGATTAAGTATGAGC 283
OY 181 TTGCATGAGGACCACTATCATATTAATGTAAGTCTTATGACCTATACATCAG 240
Db 284 CTCTCATGAGAGACCATTAATATCTAATTAATGCAAGGCTCCCTTATGATCCATCAG 343
OY 241 TGAAGATTAATCAATCAAGAACTAAAGCTAAAGATGAGATTAATGTTAATGA 300
Db 344 TGAAGACTCTCATGAAGATCGAATTAATCACTTGAAGATTCAGACATTCGAATGA 403
OY 301 GGTCAAGGTTGATGTATCAAGGTAGATGAAAATTAATTAATGTTTACCTTAAGATGC 360
Db 404 AATCAAGGTTGATGTATCAATTAAGTAAACGGTAATTAATTAATGTTTACCTTAAGATGC 463
OY 361 TGGCCACGGCGATTAAGTCCGTACAAAAGGAAATTAATTCAGACAAAACAAAGCATAG 420
Db 464 AGCTCATGCGATTAATGCTGTAACAAAAGAAATCAATCGGCAAAAACAAAGCATAG 523
OY 421 TCAACATCGTGAAGGTGAGACTCCAAAGAAAGATGCTGTTGCTTGGCAGTTGGCA 480
Db 524 TCAAGATCGTGAAGGTGAGACTCCAAAGAAAGATGCTGTTGCTTGGCAGTTGGCA 583
OY 481 AGAAGCTATATCAAGATGATGTTATTAATCTTAATGCTTCTGATATCATAGAGATAC 540
Db 584 GGACGCTACACACAGATGATGTTATTAATCTTAATGATCATGATCATGATCATGAGATAC 643

```



QY	541	TGTCATCTTATATCGTTCTCTCAGATCATTTACATTACATTCTTAAGATGAGTT	600
DB	644	GGGCGCATCTCATATTCGTTCTCTCAGAGATCATTTACATTACATTCTTAAGATGAGTT	703
QY	601	ATCACCTGACGGATTGGCTGTCGAGAAGCTTCATCTGGCTCGAGAAATCTGTCAA	660
DB	704	ATCACCTGACGAGTTGGCTGTCGAGAAGCTTCATCTGGCTCGAGAAATCTGTCAA	763
QY	661	TTCAAGAACTCATCGCCGACAAAATAGCATTAACACTTCAAGAACAACTGGGTACTTTC	720
DB	764	TTTAGAACTCATCGCCGACAAAATAGCGATTAACACTTCAAGAACAACTGGGTACTTTC	823
QY	721	TGTAAGCAATCCAGSAGACTACAAATACTAAACAAGCACACAGCAACATCTAACATCA	780
DB	824	TGTAAGCAATCCAGSAGACTACAAATACTAAACAAGCACACAGCAACATCTAACATCA	883
QY	781	AGCAAGTCAAAGTATGATGACATGTGATGCTCTTGAAACACCTCTACAACTGCTTTGAG	840
DB	884	AGCAAGTCAAAGTATGATGACATGTGATGCTCTTGAAACACCTCTACAACTGCTTTGAG	943
QY	841	TCACGACATGTAGATTCGATGGCTTGTCTTTGATCAGCACAAAATCAACAAGTCGAA	900
DB	944	TCACGACATGTAGATTCGATGGCTTGTCTTTGATCAGCACAAAATCAACAAGTCGAA	1003
QY	901	AGCTGAGGTTTGACAGTGCACAGCGAGATTCATTAACACTTCATCCCTACTCTCAAT	960
DB	1004	CGCCGAGGTTTGACAGTGCCTCAAGTAAACCATTTACCACTTTATCTCTTAAGAACAA	1063
QY	961	GTCGTAAATGGAGAAAGAAATCGCTGCTATTAATCCCTTCGTTATCGTTCAACCATG	1020
DB	1064	GTCGTAAATGGAGAAAGAAATGCTGCTATTAATCCCTTCGTTATCGTTCAACCATG	1122
QY	1021	GGTACCAATTCAAAGCCAGAACACCAAGTCCACAACGAGCTCGGAACCTAGTCCAG	1080
DB	1124	GGTACCAATTCAAAGCCAGAACACCAAGTCCACAACGAGCTCGGAACCTAGTCCAG	1183
QY	1081	CCCGCAACCTGCACCAATCTTAAATAGACTCA-----ATTCTTCTTGGT	1128
DB	1184	TCCCGCAACCTGCACCAATCTTCAACAGCTCCACAGCATTCATTTAGAGAACAAATGGT	1243
QY	1129	TAGTCAGCTGGTAGAAAAGTTGGGGAGGATATGATTTGGAAGAAAAGGCGATCTCG	1188
DB	1244	CAAAAGCACTGTTGSAAGAGAGCGATGTTATGTTCTTTGAGGAGAACATGGATTTCTG	1303
QY	1189	TTATGCTCTTCCGAAAGATTTACCATGTGAACCTGTTAAAAATCTTGAAGCAACTATC	1248
DB	1304	TTATATCCACCAAGATCTTTTCAGCAAAACAGCAGAGGATGATGAGCAAACTGCG	1363
QY	1249	AAAACACAGAGGTTTTCACACACTTAACTGCTTAAAAAAAATGTTGCTCCGCGGA	1308
DB	1364	CAGCAGGAAAGTTATCTCATAGCTAGSAGCTTAAAGAAAACGACTCCCATCTAGTGA	1423
QY	1309	CCAGAAATTTATGATTAAGCATATAATCTGTTAACTGAGGCTCATTAAGCCCTTGTGN	1368
DB	1424	TGCGAAATTTTACATTAAGCTTATGACTTACTAGCAAGAAATTCACCAAGATTTACTTGA	1483
QY	1369	AAATTAAGGTCGTATTTCTGATTTCCAGCCTTAGACAATATTAAGAACGTTGAATGA	1428
DB	1484	TAAATTAAGTGTGACAAAGTTATTTTGGGCTTTGGATTAACCTGTTGGAAACGACTAAGA	1543
QY	1429	TGAATCGCATTAATAAAAAAATTGGTAGATGTTAATTGGACTTCCTGACCAATTAAC	1488
DB	1604	TCATCTCAGAAAGCTTTAGGAAAAACCAAAATGGCAAAATTAACCTACAGATGATGATCTCA	1663
QY	1549	TATTTGCTCATTAAGCTATTAAGTATACAGAGTGAAGTGTGTACATTTTGAATGAACATGA	1608
DB	1664	AGTAGCCAAAGTTGGCAGCAAGTATACACAGAAAGCGGTTATATCTTTGATCTCTCGTGA	1723

QY	1609	TATATACTAGTATAGAGACATGCTATATGTAAGCCTCATATGGGCGCATATGACTGGAT	1668
Db	1724	TATAACCGAGTATAGGGGGATGCTATATGTAACCTACATATATGACCAATAGCCACTGGAT	1783
QY	1669	TGCAAAACATATGACCTTCTGATTAAGAAAAAGTTGACGCTCAAGCCTATATCAAAAGAAA	1728
Db	1784	TAAAAAGATAGTTGTTGTCAGAGCTGAGAGAGCGGACGCCAGCCTTATGCTAAAGACAA	1813
QY	1729	AGGTATTCCTACCTCCATCTCCAGACGCGAGATGTTAAAGCAAAATCCAACTGAGATAGTGC	1788
Db	1844	AGGTTTGACCCCTCCTTCGACAGACCATCAGATTCAGGAAATATCTGAGGCAAAAAGAGC	1903
QY	1789	AGCAGCTTATTACATAGTGTGAAAGGGGAAAAACAATTTCCACTGTTCCACTCCATA	1848
Db	1904	AGAACTATCTACAAACCGCGTGAAGAGCGATAAGAAAGTGCCACTGGATCGTATCCTTTA	1963
QY	1849	TATGGTTAGAGATACAGTGCAGTGGCTTAAACCGGTAAATTTGATTATTCCTCATAGATCA	1908
Db	1964	CAATCTTAATATATCTCTAGAGCTCAAAAACGGTATGTTAATCATACCTCATTAATGACCA	2023
QY	1909	TTACCATTAATTAATAATTTGCTTGGTTTGATGATCAACATACAAAGCTCCAAATAGCTA	1968
Db	2024	TTACCATTAATCAATTAATTTGATGTTGAGTTTGCAGAAAGCCTTTATGAGGCCCTTAAGGGTA	2083
QY	1969	TACCTTGGAAATATTTGTTGCGAGATTAAGTACTACGTAGAAACACCTGACGAAACGTCC	2028
Db	2084	TACTCTTGAGATCTTTTGGCGAGCTGCAAGTCTATGTGCAACATCCAAAGAACGTCC	2143
QY	2029	ACATTCATATGATGATGGGCAATGCGACATGACATGTGTTAGGCAAGCAAGAACACACAG	2088
Db	2144	GCATTTCAGATATATGTTTGGTAAACGTACCGCACCTATTTCACAAAGAAACAAATGTGCA	2203
QY	2089	TGAAGATCCAAATTAAGAACTTCAAGCGATGAAG-----	2125
Db	2204	AGCTGATATCCAAATCAAAACGMAAACCGAGAGAAACCTCAGACAGAAAAAATCTGA	2263
QY	2126	-----	2125
Db	2264	GGAAGAAACCCCTCGAGAGAGAAACCAAAACGGAACACGAGTCTCCAAAACCAAC	2323
QY	2126	-----CCAGTAGAGAAACACCTGCTGGAGCAGAGAGTCCCTCAAGTAGAGACTGAAA	2178
Db	2324	AGAGAACCCAGAAAGAAATCCACAGAGATCAGAAACCTCAGGTCGAGACTGAAA	2383
QY	2179	AGTAGAAGCCCACTCAAAAGACAGAGATTTGCTTGGGAAAGTAAACGATTTCTAGTCT	2238
Db	2384	GTTTGAAGAAAAACTGAGAGAGCTGAGATTTACTTGGAAAAATTCAGAGATCCAAATAT	2443
QY	2239	GAAAGCCCAATGCAACAGAACTCTAGCTGTTTACGAATTAATTTGACTCTTCAAAATAT	2298
Db	2444	CAAGTCCCAATGCCAAAGAGAGCTCTCAGAGATTTAAAAATAATTTACTATTTGGGACCCA	2503
QY	2299	GGATTAACAATATGATCTCATGTCAGACAGACAGAAAAATTACTGGGTTGTTAAAAAGAGTAA	2358
Db	2504	GGACACACATATCTATTATGCGACAGAGCTGAAAAACTATTGGCTTTATTAAAGAGAGTAA	2563
RESULT 7			
ABK15103			
ID	ABK15103 standard; DNA: 2647 bp.		
XX	ABK15103;		
XX	08-MAY-2002 (first entry)		
XX	DNA encoding Streptococcus pneumoniae BVH-11.		
XX	BVH-3; BVH-11; vaccine; meningitis; otitis media; bacteraemia;		
KW	pneumonia; streptococcal bacterial infection; gene; ds.		
XX	Streptococcus pneumoniae.		
XX	Key	Location/Qualifiers	
PH			





```
Db 1424 TCGAGAAATTTACAAATAGGCTTATGCTTACTACTACGACAAAGATTCACCAAGATTACTTGA 1483
Qy 1369 AAATTAAGGTCGTAAATCTGATTTTCCAAAGCCTTAGACAAATTTATTAGACGCTTGAATGA 1428
Db 1484 TAATTAAGTCGACAAAGTGTATTTTGAGGCTTTGGATTAACCTGTTGGAAGACGACTCAAGGA 1543
Qy 1429 TGAATGCACCTAAATTAAGAAATTTGGTAGATGATTATTTGGCATTTCTAGCACCATTAC 1488
Db 1544 TGTCTCAAGGATAAAGTCAAGTAGTGATGATATTTCTGCTTCTTACTCTCGATTCG 1603
Qy 1489 CCATCCAGACGACTTGGCAACCAATTTCAAAATGATATACCTTAAGACGAACTTCG 1548
Db 1604 TCATCCGAAGCCTTAGGAAAACCAATTCGCAATTTACCTACTACTATATGATGATTTCA 1663
Qy 1549 TATTTGCTCAATTAAGCTGATTAAGTATACACGCTGATGTTTATCTTTTATGAATGA 1608
Db 1664 AGTAGCCCAATGGGACGCAAGTACACAGAAAGAGGTTTATATCTTTGATCTCTGTGA 1723
Qy 1609 TATATACGATGATGAAGACATGATATTAAGCCTCATATGGCCATATGCTAGTGGAT 1668
Db 1724 TATATCCAGATGATGAGGGGATGCTTATGCTTACCTACATATGACCATAGCCATGAT 1783
Qy 1669 TGGAAAGATAGCCTTCTGATTAAGAAAAAGTTGCAGCTCAAGCCTATCTAAAGAAA 1728
Db 1784 TAAAAAGATAGTTTGTCTGAGCTGAGAGAGGCGCCAGCCTTATGCTTAAAGGAA 1843
Qy 1729 AGGTATCTCACTCCATCTCCAGACGAGATGTTAAAGCAAAATCCAACTGAGATAGTGC 1788
Db 1844 AGGTTTGAACCCCTTCCTTGCAGACGATCAGATTCAGAAATTAAGTGAAGCAAAAGAGC 1903
Qy 1789 AGCAGCTATTTACAAATGCTGTGAAGGGGAAAAACGAATTCACCTGCTTGACTTCATA 1848
Db 1904 AGAAGCTATCTACACGCGCTGAAAGAGAGCTAAGAAAGTCCACTGATGCTTATGCTTGA 1963
Qy 1849 TATGTTGACATACAGTTAGGTTTAAAAACGGTAATTTGATTTATTCCTCATAGAGTCA 1908
Db 1964 CAATCTTCAATATACGTGTAGAAAGTCAAAAAAGGTAATTAATATCTATCTATTAATGACA 2023
Qy 1909 TTACATATATTTAAATTTGCTTGTGATGATCAACATACAAAAGCTTCAAAATGGCTA 1968
Db 2024 TTACATATATCAAAATTTAGTGTGACGAAGGCGCTTATGAGCACCCTTAAGGGTGA 2083
Qy 1969 TACCTTGAAGATTTTGTTCGACGATTAAGTACTAGTGAACACCCCTGACAGAGTCC 2028
Db 2084 TACTCTGAGATCTTTTGGCGACTGTCAGTACTATGTCGAACATCCAAACAGCTCC 2143
Qy 2029 ACATCTTAATGATGATGGGCAATGCCAGTGAAGCATGTTGTTAGCAAGAAGACACAG 2088
Db 2144 GCATTCAGATTAATGCTTTGGTAAAGCTAGCCAGACATGTTCAAAAGAAACAAAATGCTCA 2203
Qy 2089 TGAAGATCCAAATTAAGAACTTCAAAGCGGATGAAG----- 2125
Db 2204 AGCTGATACCAATCAAAACGAAAAACGAGAGAAACCTCAGACAGAAACAACTTGA 2263
Qy 2126 ----- 2125
Db 2264 GGAAGAAACCCCTCGAAGAGAGAAACCAAAACGAGAAACAGAGTCTCCAAAAACCAAC 2323
Qy 2126 -----CCAGTAGAGAAACACCTGCTGAGCCAGAAAGTCCCTCAAGTAGAGACTGAAA 2178
Db 2324 AGAGGAACCAAGAAAGATCACACAGAGATCAGAAAGAACTCAGAGCTGAGACTAAAA 2383
Qy 2179 AGTAGAAGCCCACTCAAGAAAGCAGAGTTTGTTCGGAAGTAAGGATTTCTAGTCT 2238
Db 2384 GGTGAAGAAAAAAGTGAAGAGCTGAAGATTTACTTTGAAAAAATCCAGATCCAAATTAAT 2443
Qy 2239 GAAAGCCATGACACAGAAACCTTACTGTTTACGAATTAATTTGACCTCTTCAATTTAT 2298
Db 2444 CAAGTCAATGCGCAAGAGACTCTTCACAGGATTTAAAAAATTAATTTCTATTGCGACCCA 2503
Qy 2299 GGATTAATAGTATGATGACAGAGCAAGAAAAATTAATTTACTGTTTAAAGAGATTA 2358
Db 2504 GGACAAACATACATATATGACAGAGCTGAAAAACCTAATTTGGCTTTATTAAAGAGAGTAA 2563
```

```
RESULT 8
ID AAA08557
XX AAA08557 standard; DNA; 2478 BP.
AC AAA08557;
XX
DT 19-JUL-2000 (first entry)
XX
DE S. pneumoniae 92 kDa human C3-degrading protein coding sequence.
XX
KW Human C3-degrading protein; 92 kDa; immunostimulatory; vaccine;
KW inhibitor; inflammation; organ rejection; xenotransplantation; ss.
XX
OS Streptococcus pneumoniae.
XX
PN MO200017370-A1.
XX
PD 30-MAR-2000.
XX
PF 24-SEP-1999; 99WO-US22362.
XX
PR 24-SEP-1998; 98US-0101736.
XX
PR 31-MAR-1999; 99US-0283094.
XX
PA (MING ) UNIV MINNESOTA.
XX (AMCY ) AMERICAN CYANAMID CO.
XX
PI Hostetter MK, Finkel DJ, Cheng Q, Green BA, Masi AW;
XX
XX MPI: 2000-283594/24.
XX P-PSDB; AAY91939.
XX
PT Isolated polypeptide is used to stimulate immune system and immunize or
PT treat a mammalian subject against Streptococcus pneumoniae infection or
PT colonization
XX
XX Claim 1, Page 55-57; 63pp; English.
XX
CC The present sequence, isolated from Streptococcus pneumoniae, encodes a
CC human C3-degrading protein (see AAY91939) of about 92 kDa. This
CC sequence may encompass a smaller 20 kDa polypeptide coding sequence
CC (AA08556) also having human C3-degrading activity.
CC The DNA sequences can be used for producing an immune response to
CC Streptococcus pneumoniae in a mammal. Antibodies against the proteins
CC can be used to inhibit S. pneumoniae-mediated C3 degradation.
CC C3-mediated inflammation and rejection in xenotransplantation can be
CC inhibited by expressing the nucleic acid sequences on the surface of an
CC organ of an animal. In particular, the polypeptides are useful for
CC stimulating the immune system and are effective to immunize or treat a
CC mammalian subject against Streptococcus pneumoniae infection or
CC colonization.
XX
SQ Sequence 2478 BP; 837 A; 510 C; 535 G; 596 T; 0 other;
XX
Query Match 42.4%; Score 1011.8; DB 21; Length 2478;
Best Local Similarity 66.2%; Pred. No. 3.8e-244;
Matches 1628; Conservative 0; Mismatches 688; Indels 144; Gaps 6;
```

Db 240 GACCTTCATGAGACCATATCATTTACTATATGCGAAGCTCCCTTATATGCAATCAT 299  
QY 238 CACTGGAAGATTTACTCATGGAAGATCCAACTATAGCTAAAGATGAGATATGTTAA 297  
Db 300 CACTGGAAGATTTACTCATGGAAGATCCAACTATAGCTAAAGATGAGATATGTTAA 359  
QY 298 TGAGGTCAAGGATGATATGTTATCAAGTATGGAAGATTAATGTTTACCTTAAGGA 357  
Db 360 TGAATCCAAAGGATGTTATTTATCAAGTATGGAAGATTAATGTTTACCTTAAGGA 419  
QY 358 TCGTCCCAAGGATGATATGTTATCAAGTATGGAAGATTAATGTTTACCTTAAGGA 417  
Db 420 TCGAGCTCATGCGATATATTTGAGCAAAAGAGATTAATGTTTACCTTAAGGA 479  
QY 418 TACTCAACATGCTGAAGGTGATGCTCAAGAAAGATGCTGCTTCCCTTGACACTTC 477  
Db 480 CACTCATATACGAGGAGGATGCTTCT-----AACGATCAAGCATGATGTTGACGCAAGC 533  
QY 478 GCAAGGACGCTATACAGATGATGTTATATCTTTAATGCTTCTGATATCATAGAGA 537  
Db 534 CCAAGGACGCTATACAGATGATGTTATATCTTCAATGCAATCTGATATCATAGAGA 593  
QY 538 TACTGATGATGTTATATGCTTCTCTCATGAGATCATTTACCATTTACCTTAAGGA 597  
Db 594 CACGGGATGATGTTATATGCTTCTCTCATGAGATCATTTACCATTTACCTTAAGGA 653  
QY 598 GTTATACGATGAGGATGTTGCTGTCAGAAAGCTTCTCTATGCTGTCAGAAATCTGTC 657  
Db 654 GTTATACGATGAGGATGTTGCTGTCAGAAAGCTTCTCTATGGAATGG-----698  
QY 658 AATTCGAAGATCTATGCGCGACAAAATAGCATTAACATCTTCAAGAACTGGGTACC 717  
Db 699 -GAAGAGGATGCTCGTCTTCTTCAAGTCTAGTTATATGCAAACTCAGCTCAACAA 757  
QY 718 TTCTGTACGAATCCAGGAATCAAACTAATAACACAGAACACAACTAAG 777  
Db 758 GATTGTCCAGGAACCAACATCT-----GACTGTCACTCCCACTTA 797  
QY 778 TCAAGCAAGTCAAGATTAAGATGATGATCTTGAAGACGTCACAACTGCTT 837  
Db 798 TCAATCAAAATCAAGGAGAAACATTTCAAGCTTTTACGAAATGATCTTAACCTT 857  
QY 838 GAGTCAACGATATGATATCTGATGCTTGTCTTGTATGATCCAGCAAACTCAAGTCC 897  
Db 858 ATCAGAAAGCATGATGATCTGATGCTTATTTGACCCAGGCAAACTCAAGTCC 917  
QY 898 AACAGCTAGAGTGTTCAGATGCTCAGACAGATATTAACCTTCACTTCCCTTACTCA 957  
Db 918 AACCCGAGAGGTGATCTCTCCCTCATGTTAACCATTAACCTTATGCACTTATGACA 977  
QY 958 AATGCTGATTTGGAAGAAAGATGCTGATATATCCCTTGTATGTTTAAACA 1017  
Db 978 AATGCTGATTTGGAAGAAAGATGCTGATATATCCCTTGTATGTTTAAACA 1037  
QY 1018 TTGAGTACAGATTCAGAGCCAGAAACAGATCCAAACGACTCCGAACTAGTCC 1077  
Db 1038 TTGAGTACAGATTCAGAGCCAGAAACAGATCCAAACGACTCCGAACTAGTCC 1097  
QY 1078 AGGCCCGCAACCTGACCAATCTTAAATAGACTCA-----ATTCTTCTTT 1125  
Db 1098 AATCTCCCAACCTGACCAATCTTAAACAGCTCCCAAGCAATCCAAATGGAATTT 1157  
QY 1126 GGTATGACCTGTTAGAAAGTTGGGAGAGATATGTTTGAAGAAAGGACATCTC 1185  
Db 1158 GGTCAAGAAAGCTGTTGAAAGATGAGCGATGTTATGCTTTGAGGAGATGAGATTTC 1217  
QY 1186 TCGTTATGCTTTGCGAAGATTTACCATCTGAAGCTTTAAATCTTGAAGCAAGTT 1245  
Db 1218 TCGTTATATCCAGCCAAAGATCTTTACAGAGAAACAGACGAGCATGATACCAACT 1277  
QY 1246 ATCAAAACAGAGAGCTTTTACACACTTTTAACTGCTTAAAAAGAAATGTGCTCTCG 1305  
Db 1278 GGGCCAGACGAAAGTTTATCTCAATAAGCTAGAGACTAAGAAATGACCTCCATCTAG 1337  
QY 1306 TGACCAAGAAATTTATGATAAAGCATATATCTGTTAAGTACGAGCTCATAAAGCTTGT 1365  
Db 1338 TGATCGAAGATTTTACATTAAGCTTTTACCTTACTGACCAAGATTTACCAAGTTTACT 1397  
QY 1366 TGNAAATGAAGGCTGTAATCTGATTTTCCAAAGCTTGAACAAATTTATAGAACGCTTGA 1425  
Db 1398 TGATTAATAAGGCTGACAAAGTGTATTTGAGGCTTTGGATTAACCTGTTGGAAGCACTCA 1457  
QY 1426 TGATGAATCGACTATATAAGAAATTTGATGATGATTTATGSCATTTCTAGCACTAAT 1485  
Db 1458 GGATGTCCCAAGTGTAAAGTCAAGTTAGTGATGATATTTCTTCCCTTCTAGCTCCGAT 1517  
QY 1486 TACCATCCAGAGGACTTGGCAAAACCAATTCCTCAATGAGTATCTGAGAGCAAGT 1545  
Db 1518 TCGTATCCAGAACCTTTAGCAAAACCAATTCGCAATTTACTTACCTGATGATGATGAT 1577  
QY 1546 TCGTATTTGCTCAATTTAAGCTATTAAGTATTAACACGCTCAGATGTTTATGATGAACA 1605  
Db 1578 TCAAGTACCAAGTTGGCAGGCAAGTACACACAGAAAGAGGTTATATCTTGTATCTCG 1637  
QY 1606 TGATATATACGATGATGAAGAGATGCAATATGTAAGCTCATATGAGCCATAGTCACTG 1665  
Db 1638 TGATATACCAAGTATGAGGAGGATGCTTATGTAATCTCCATGACCATAGCCACTG 1697  
QY 1666 GATTGGAAGATATGACCTTCTGATTAAGAAAGTTGCAAGCTCAACCTATACTAAGA 1725  
Db 1698 GATTAAAGAAATGATGTTGTTGTAAGCTGAGAGGCGGACCCAGCTTATGCTTAAGA 1757  
QY 1726 AAAAGTATCTTACCTCATCTCCAGCAGAGATGTTAAAGCAATCCAACTGAGATAG 1785  
Db 1758 GAAAGTTTGACCCCTCTCCAGACACATCAGATTTAGGAATATCTGAGCAAAAG 1817  
QY 1786 TGACAGCTATTTTCAATGCTGTGAAGAGGAAACAAATTCCTGCTGATCTCC 1845  
Db 1818 AGCAGAACTATCTACCAACCGCGGAAAGCAAGCTTAAGAGTGCACCTTATCTATGCG 1877  
QY 1846 ATATATGTTGAGCATCAGTGAAGTTAAAGAGTAAATTTGATTTTCTCTAATAAGA 1905  
Db 1878 TTACATCTTCAATATATCTGATAGATCAAAACAGTATGTTTATCAATCTCAATTTGA 1937  
QY 1906 TCAATACCATTAATTTAAATTTGCTTGTGATGATCACACATCAAAAGCTCCAAATGG 1965  
Db 1938 CCATTAACATTAATCAATTAATTTGAGTGTGTTGACGAAGGCTTATAGGACCTTAAGGG 1997  
QY 1966 CTATACCTTGAAGATTTGTTTGGAGAGATTAAGTACTAGTAAACACCTTGACGAGC 2025  
Db 1998 GTATACCTTGTAGAGATCTTTTGGCGACTGTCMACTATGTCMACTCCAAACGAGC 2057  
QY 2026 TCCCATTTCAATGATGATGAGGCAATGCCAGTGAAGATGTTAGGCAAAAGAGCA 2085  
Db 2058 TCCGATTCAGATATATGTTTGTGTAACGCTAGGACCAATGTTCAAGAAACAAATATGG 2117  
QY 2086 CAGTGAATTCCAATTAAGACTTCAAGCGGATGA-----2121  
Db 2118 TCAAGCTGATATCAATCAAAACGAAACCAAGAGAGAGAAACCTCAGACAGAAAC 2177  
QY 2122 -----2121  
Db 2178 TGAGGAAGAAACCCCTGAGAGAGAAACCGCAAGGAGAAACAGAGTCTCCAAAC 2237  
QY 2122 ---AGACCCAGTATAGGAAACACCTGCTGAGCCAGAGATCCCTCAAGTATGAGCTGAAA 2178  
Db 2238 AACAGAGAAACCAAGAAATCACAGAGATCAAGAAAGCTTACGATCGAGACTGAAAA 2237  
QY 2179 AGTAGAAGCCCACTCAAGAGAGAGAGATTTGCTTGGCAAGTAAAGGATTTAGTCT 2238  
Db 2298 GGTGGAAGAAAACTGAGAGAGGCTGAAGATTTACTTGGAAATATCAAGATCCAAATAT 2357  
QY 2239 GAAAGCCATATGCAACAACTAGTGTGTTTACGAATATATGACTCTTCAATATAT 2298  
Db 2358 CAACTCCAAATGCCAAAGAGACTCTCACAGATTTAAAAAATATTTACTATTTGGCACCCA 2417



Db 1218 TCCTATATCCAGCCAGAGATCTTTCAGCAGAAACAGCAGCGATTTAGTAACT 1217  
 QY 1246 ATCAAAACAGAGAGTGTTCACACACTTACGTGCTAAAAAGAAATGTGCTCG 1305  
 Db 1278 GGCCAAAGAGAAAGTTTATCTATAGCTAGAGACTAAGAAACAGACTCCCATCTAG 1337  
 QY 1306 TGACCAAGATTTTATGATTAAGCATTAATCTGTTAACTGAGGCTCATTAAGGCTTGT 1365  
 Db 1338 TGATCGAGATTTTACAAATAAGGCTTATGACTACTAGCAAGATTCACCAAGATTACT 1397  
 QY 1366 TGAATTAAGGCTGCTAATCTGATTTCCAGGCTTAGAACAAATATTAAAGCTTGAA 1425  
 Db 1398 TGATTAATTAAGGCTGCAACAGTATTTTGGAGCTTGGATTAACCTGTTGAAGCAGCAA 1457  
 QY 1426 TGATGAATGCACTAATTAAGAAATAATGTAGATGATTTTGGCATTTCTAGACCAAT 1485  
 Db 1458 GGATGTCGCCAAGTATGAAGTCAAGTTAGTGAATATTCTTGCTCTTACGCTCCGAT 1517  
 QY 1486 TACCATCCAGAGCGACTTGGCAACCAATTTCTCAAAATTGATATCTGAAGCAGAGT 1545  
 Db 1518 TCGTCATCCAGAACGTTTAGAAACCAATGCGCAAAATTACCTACACTGATGATGAT 1577  
 QY 1546 TCGTATTTGCCAATAGCTATAGCTATAGTATACAGCTCAGATGCTTACATTTTGATGACA 1605  
 Db 1578 TCAAGTAGCCAAAGTGGCAGGCAAGTACACACAGAAAGAGGTTATATCTTGTCTCTCG 1637  
 QY 1606 TGATATATGATGATGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1665  
 Db 1638 TGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1697  
 QY 1666 GATTGGAAGAGATAGGCTTCTGATGAAGAAAGTTTGCAGCTCAAGCTTATCTAAAGA 1725  
 Db 1698 GATTAAAGAAAGATAGTGTGCTGAGCTGAGAGCGCGCAGCCAGCTTATGCTAAGA 1757  
 QY 1726 AAAAGTATCTCTACCTCATCTCCAGACGCAATGTTAAAGCAATCAACTGGAGATAG 1785  
 Db 1758 GAAAGTTTGTACCCCTTCCTCGACAGCATCGAGATTCCGAAATTACTGAGCGAAAGG 1817  
 QY 1786 TGCAGCAGCTATTTTACATAGTGTGAAGGGAAGAAAGCAATTCCTGCTGACTTCC 1845  
 Db 1818 AGCAGAAAGCTATCTACACGCGCTGAAGAGCACTAAGAGTGCCTTATGCTATGCC 1877  
 QY 1846 ATATATGTTGAGCATACAGTATGAGTTTAAACGTTAATTTGATTAATCTCATTAAGA 1905  
 Db 1878 TTACAAATCTCATATATACCTAGAGCAAAAGGAGTTTATATATCTATCTATATGA 1937  
 QY 1906 TCATTTACATTAATTAATTTGCTGTTGTTGATGATCAGCATACAGACTCCAAATGG 1965  
 Db 1938 CCATTTACCATACATCAAAATTTAGTGGTTTGACGAAGGCTTATGAGGACCTTAAGGG 1997  
 QY 1966 CTATACCTTGGAGATTTGTTTGGCAGATTAAGTACTAGTAAGACACCTGACGAGG 2025  
 Db 1998 GTATACCTCTTGAAGATCTTTTGGCAGCTGCAAGTACTATGTAACATCAAGCAAGG 2057  
 QY 2026 TCCACATTTCTAATGATGATGGGCAATGCCAATGAGCATGTTTGAAGCAAGAACCA 2085  
 Db 2058 TCCGCAATTCAGATTAATGTTTGTGTAACGCTACGACCAATGTTCAAGAAACCAAAATGG 2117  
 QY 2086 CACTGAAGATCCAAATPAGAACTTCAAGCGAGTA----- 2121  
 Db 2118 TCAAGCTGATACCAATCAAAAGCAAAACCAAGCAGAGAGAAACCTCAGACAAAAACC 2177  
 QY 2122 ----- 2121  
 Db 2178 TGAGGAAGAAACCCCTCGAGAGAGAAACCGCAAGCGAGAAACAGAGTCTCCAAAACC 2237  
 QY 2122 ---TGAGCCAGTAGAGAAACACCTCTGAGCCAGAAAGTCCCTCAAGTAGAGACTGAAAA 2178  
 Db 2238 AACAGAGGAACAGAGAAATCACACAGAGATCAGAAAGAACCTCAGTGCAGACTGAAAA 2297  
 QY 2179 AGTAGAAGCCCAACTCAAGAGACAGAGATTTTGTGCTTGCAGAAAGTAACGAGATTCTACTCT 2238  
 Db 2298 GGTGTAAGAAAGAACTGAGAGAGGCTGAGATTTACTTGAAGAAATCCAGAGATCAATTAT 2357

QY 2239 GAAAGCCATGCAACAGAACTCTAGCTGTTTACGAATTAATTTGACTCTCAATTAT 2298  
 Db 2258 CAACTCCAAATGCCAAGAGACTCTCACAGATTTAAAAAATTAATTTACTATTTGGCACCCA 2417  
 QY 2299 GGATTAACAAATAGTATCATGTGCGAGAGCAAAAAATTAATTCGCTGTTTAAAGGAACTAA 2358  
 Db 2418 GGACAAACAAATACTATTTATGTCAGAGAACTGAAAAAATCTGCTTATTAAGAGAGACTAA 2477  
 RESULT 10  
 AAA47605 standard; DNA; 2531 BP.  
 ID AAA47605:  
 AC  
 XX  
 DT 20-OCT-2000 (first entry)  
 XX  
 DE Recombinant variant of Sp36 gene (Sp36B) of *S. pneumoniae*.  
 XX  
 KW Streptococcus pneumoniae; infection; vaccine; coiled coil region;  
 KW histidine triad residue; Sp36; antibody; otitis media;  
 KW nasopharyngeal infection; bronchial infection; bronchitis; sepsis;  
 KW meningitis; lobar pneumonia; ds.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 FH key Location/Qualifiers  
 FT CDS 1..2531  
 FT /tag- a  
 FT /product= sp36B polypeptide  
 PN W0200037105-A2.  
 XX  
 PD 29-JUN-2000.  
 XX  
 PF 21-DEC-1999; 99MO-US30390.  
 XX  
 PR 21-DEC-1998; 98US-0113048.  
 XX  
 PA (MED1-) MEDIMUNE INC.  
 XX  
 PI Johnson LS, Koenig S, Adamou JE;  
 DR WPI: 2000-452129/39.  
 DR P-PSDB; AAB01469.  
 XX  
 PT Vaccine useful for prophylaxis and treatment of pneumococcal infections  
 PT such as otitis media, nasopharyngeal and bronchial infections,  
 PT comprises Streptococcus pneumoniae proteins  
 XX  
 PS Disclosure; Page 69-70; 70pp; English.  
 XX  
 CC Although a number of proteins have been suggested as being involved  
 CC in the pathogenicity of Streptococcus pneumoniae, there still remains  
 CC a need to identify polypeptides having epitopes in common from  
 CC various strains of *S. pneumoniae* in order to utilize such  
 CC polypeptides in vaccines to protect against a wide variety of  
 CC *S. pneumoniae*. New vaccine compositions are described which comprise a  
 CC Streptococcus pneumoniae polypeptide (or fragments) of 80 - 680 amino  
 CC acids in length that comprise at least one histidine triad residue  
 CC (HxxHxH) or a coiled-coil region, or an antibody directed against  
 CC these features. The vaccine is useful in protecting against infection  
 CC by Streptococcus pneumoniae. The vaccine composition comprising  
 CC antibodies to is useful for passive immunization for treating  
 CC pneumococcal infections which includes otitis media, nasopharyngeal  
 CC and bronchial infections.  
 CC  
 SO Sequence 2531 BP; 861 A; 508 C; 550 G; 609 T; 3 other;  
 Query Match 41.5%; Score 991; DB 21; Length 2531;  
 Best Local Similarity 65.7%; Pred. No. 6,6e-239;  
 Matches 1609; Conservative 1; Mismatches 697; Indels 141; Gaps 6;









Dh	1918	ANCCGACGAGGTCTACCTCTCCCTCATGTAACGATTACACCTTATCCCTTATGAAACA	977
Qy	958	AATGCTGAATTTGGAGAAGAAATCGCTGTATTATTCCCTTGCTATCGTTCAAAACA	10117
Dh	978	AATGCTGAATTTGGAAAAAGAAATTCCTGTATTATTCCTTCGTATTGTCGTAACCA	1037
Qy	1018	TTGGGTACCAATTTCAAGGCCAACAACCAAGTCCACACCGCATCCGGAAACCTAGCC	10777
Dh	1038	TTGGGTACCAATTTCAAGACCACAACCAACCAATCCCAATGACATCCGGAACTAGCC	1097
Qy	1078	AGGCCCGCAACCTCGCACCAATCTGTTAAATAGACTCAAAATCTTC-----TTC	1125
Dh	1098	AAGTCCGAACCTCGACCAATCTCTCACACCGCTCCAAACATTCATTTGATGAAATTT	1157
Qy	1126	GGTATGTCAGCTGTACGAAAAGTTGGGGAAGGATATGATTTCGAAGAAAAGGCGATTC	1185
Dh	1158	GGTCAAAAGAAACCTGTTGAAAAAGTAGCGCGATGTTATGCTTTGAGGAGAAATGCGAGTTTC	1217
Qy	1186	TCGTTATGCTTTTCCGAACAATTTACCATCTGAACCTGTTAAAAATCTTTGAAAGCAAT	1245
Dh	1218	TCGTTATATCCACACCAGAGATCTTTTCAGCAGAAACGAGAGAGCATGTAGAGAACT	1277
Qy	1246	ATCAAAACAGAGAGCTTTTCACACACTTTTAATGCTTAAAAAAGAAATGTTGCCTCCG	1305
Dh	1278	GCCCAAGCAGAAAGTTTATCTATAGCTAGGAGCTTAAGAAAACCTGACCTCCCACTAG	1337
Qy	1306	TGACCAAGAAATTTTATGATTAAGCATATATCTGTTTACTGAGGCTCATAAAGCCTGTT	1365
Dh	1338	TGATCGAAGAAATTTTACAAATTAAGGCTTATGACTTACGACAGAAATTCACCAAGATTACT	1397
Qy	1366	TGMAATTAAGGTCGTAAATTCGATTTCCACAGCTTATGCAAAATTTATGAAAGCTTGAA	1425
Dh	1398	TGATTAATTAAGCTCGACAAATGATTTTGGCTTTGGATTAACCTGTTGGAACGACCTAA	1457
Qy	1426	TGATGAATCGACTATTAAGAAAAAATTTGGTAGATTTATGGCAATTCCTAGACCAAT	1485
Dh	1458	GGATGTCCCAAGTATTAAGTCAAGTTAGTGGATGATATTTCCCTTCTTAGCTCCGAT	1517
Qy	1486	TACCATCCAGAGGACTTGGCAAAACCAATTCCTAAATTTGAGTATACTGAAAGCAAGT	1545
Dh	1518	TCGTCATCCAAAGCTTTAGGAAAAACCAATGGCCAAATTTACTCTACGTAGATGAAT	1577
Qy	1546	TCGATTTGCTCAATTAAGCTATTAAGTATACACGTCAGATGTTACATTTTGATGAACA	1605
Dh	1578	TCAGTATGCCCCAAGTTGGCAGCAAGTACAAACAGAAAGCGTTATATCTTTGATCTCG	1637
Qy	1606	TGATATATACGTATTAAGAGAGATGCAATATGTAACGCTCATATGGCCATAGTCACTG	1665
Dh	1638	TGATATATACCAATGATAGGGGAGTGCCTTATGTAATCTCACATATGACCATAGCCACTG	1697
Qy	1666	GATTGAAAAAGATAGCCTTTCTGATTAAGGAAAAAGTTGCAGTCAAGCCTTACTAAAGA	1725
Dh	1698	GATTAAAAAAGATAGTTGTCTGTAAGCTGAGAGCGGAGCCAGCGCTTATGCTAAAGA	1757
Qy	1726	AAAAAGTATCTTACCTTCATCTCCAGACGAGATGTTAAAGCAATTCACACTGAGATAG	1785
Dh	1758	GAAAGGTTTGACCCCTTCCTTCGACAGACATCAGGATTCAGGAAATACTGAGGCAAAAG	1817
Qy	1786	TGCGACGACTATTACAAATGCTGTAAGGGGAAAAACCAATTCACAGCTTTCAGCTTCC	1845
Dh	1818	AGCGAAAGCTATCTACAAACGCGGTGAAGAGAGCTTAAGAGGTCGATGATGATGCC	1877
Qy	1846	ATATATGGTTAGCATATCAGTTCAGGTTAAAAACGTAATTTGATTTATTCCTCAATAGA	1905
Dh	1878	TTACAATCTTCAATATCTGTAAAGTCAAAAAAGGTATGATCATATCCTCATTTATGA	1937
Qy	1906	TCATATACCAATATTTAAATTTGCTTGGTTGTGATCAACATCAAAAGCTCCAAATGG	1965
Dh	1938	CCATTACCAATCAATCAAAATTTGAGTGGTTTGACGAAGCCCTTTATGAGGCAACCTAAGG	1997
Qy	1966	CTATATCCTTGGAGATTTGTTTGCACGATTAAGTACTACGTAGAACAACCTGACGAACG	2025
Dh	1998	GTAATATCTTGGAGATCTTTTGGCGACTGTCAAAGTCTATGTCGAACATCCAAACGAACG	2057

QY	2026	TCACATTC	TATATGATGATGGGCAATGCCAGTGAGCATGTGTTAGCCAGAAAGACCA	2085
Db	2058			
QY	2086	CAGTCAACATCCAAATTAAGACTTC	CAAGCGGATGAAGACGACAGTAAACACCTGC	2145
Db	2118	CCAGACAGCTAAACCTGATGAAAGATTAAGAACATCATATTAAGTAACTGAGCCAACTACACC	2177	
QY	2146	TCAGCCAGAAAG	2156	
Db	2178	TGAATCTGATG	2188	
RESULT 12				
AAV27356				
ID	AAV27356	standard; DNA; 2290	BP.	
XX	AAV27356;			
AC	AAV27356;			
XX	02-OCT-1998	(first entry)		
DT				
XX	Streptococcus pneumoniae	SP0042	nucleotide.	
DE				
XX	Streptococcus pneumoniae;	antigen; vaccine; infection;	diagnosis;	
KW	detection; pneumonia; otitis media; meningitis; ss.			
XX	Streptococcus pneumoniae.			
OS				
XX	Key	Location/Qualifiers		
FT	CDS	2..2290		
FT	/*tag=	a		
FT	/product=	"SP0042"		
FT	/transl_except=	(pos:152..154,aa:Xaa)		
FT	/transl_except=	(pos:1406..1408,aa:Xaa)		
FT	/transl_except=	(pos:1430..1432,aa:Xaa)		
FT	/note=	"no stop codon given; Xaa is unspecified"		
XX	W09818930-A2.			
PN				
XX	07-MAY-1998.			
PD				
XX	30-OCT-1997;	97WO-US19422.		
PF				
XX	31-OCT-1996;	96US-0029960.		
PR				
XX	(HUMA-) HUMAN GENOME SCI INC.			
PA				
XX	Choi GH, Hiromocky J A, Johnson LS, Kunsch CA;			
PI				
XX	WPI: 1998-272224/24.			
DR	P-PSDB: AAM55095.			
XX				
PT	Nucleic acid encoding antigenic peptide(s) from Streptococcus			
PT	pneumoniae - or their epitope-containing fragments, useful in			
PT	protective or therapeutic vaccines, and for diagnosis			
XX				
PS	Claim 1; Page 61-62; 118pp: English.			
XX				
CC	The present sequence encodes a protein from Streptococcus pneumoniae.			
CC	The nucleic acid sequence encoding the Streptococcus pneumoniae protein			
CC	can be useful in vaccines for inducing protective antibodies against			
CC	Streptococcus pneumoniae, for treatment or prevention of infection e.g.			
CC	pneumonia, otitis media or meningitis. Probes based on the nucleic acid			
CC	are used to detect Streptococcus infection (by usual hybridisation or			
CC	amplification methods), also for isolating Streptococcus genes or their			
CC	allelic variants. The protein can be used similarly to detect specific			
CC	antibodies in standard immunoassays, especially for diagnosing or			
CC	monitoring infections. Antibodies which bind the protein are used to			
CC	detect corresponding antigens, to purify the protein and for passive			
CC	immunisation (optionally coupled to a toxin). Vaccines are administered			
CC	e.g. by injection, orally or through the skin, typically at 0.01-1000			
CC	(especially 10-300) mu g/ml per dose.			

XX Sequence 2290 BP; 766 A; 474 C; 498 G; 547 T; 5 other;

Query Match	41.3%	Score 987.6;	DB 19;	Length 2290;
Best Local Similarity	67.7%	Pred. No. 4.5e-238;		
Matches 1481; Conservative	0;	Mismatches 645;	Indels 60;	Gaps 5;

QY	1	TTCTTACAGAGTTGGGACTGTATCAACCTGTAAGCGTTAAGGAAA---TAATCGTTTTC	57
Db	4	TTCTTATGAAGCTTGGTGTGTACCAACCTGGTCAGGTTATGAAGAGTGTAAATCGAGTTTC	63
QY	58	CTATATATGATGAAAAACAAGCGACGAAAAAAGCGAATTTGATCCTCGATGAGGTAG	117
Db	64	TTTATATGATGTGTATCAGGCTGTGTCAAAAGCGAAGAACTTGACACCAGATGAAGTCAG	123
QY	118	CAAGCGTGAAGGAAATCAATGTCTGAGCAAAATCGTCATCAAGATTAACGACCAGGCTATGT	177
Db	124	TAAGAGGAGGAGGATCAACGCCGCAAAATNGTNTCAAGATTAACGATTCAGATCAAGGTTATGT	183
QY	178	CACATTCATGAGGAGACACATCATATTATTCATAGGTAGGTTCCTTATGACCGTATCAT	237
Db	184	GACCTCTCAATGAGGAGACATTATATCATTTACTATATATGSCAAGGTTCTTATGATGCCATAT	243
QY	238	CAGTGAAGAAATTAATCATGAAAGATCCAACATTAAGCTTAAAGATGAGATATTGTAA	297
Db	244	CAGTGAAGAGCTCCTCATGTAAGATCCGAATTTATCATGTTGAAGATTCAGACATTTGCA	303
QY	298	TGAGGTCAAGGGGTGATATGTTATCAAGGTAAATGGAATAATCTATGTTTACTTTAAGGA	357
Db	304	TGAATTAAGGGGTGTTATGTCTTTAAGGTAAACGGTAAATCTATGTTTACTTTAAGGA	363
QY	358	TGCTGCCAGCGCGATACAGTCCGTCGTCACAAAGAGAAATCAATCGACAAAAACAGAGCA	417
Db	364	TGCAGCTCATGCGGATATATATTGCGACAAAAGAGATTAACGTCAGAAAGCAGAGAC	423
QY	418	TAGTCAACATCTGGAAGGTGGAACCTCAAGAAACGATGCTGCTTGGCTTGGCACGTTC	477
Db	424	CAGCTCAATTAATCAT-----AACTCAAGACAGATTAATGCTGTTGTCGACCGCAGAC	474
QY	478	GCAAGGACGTTACTACAGATGATGGTATATTCCTTAATGCTTGATATCATATGAGGA	537
Db	475	CCAAGGACGTTATACCAACGATATGCGTATATCTTTCAATGCATCTGATATCATTTGAGGA	534
QY	538	TACTGTGATGCTTATATGCTTCTTCATGAGAGATCAATTAACATTCATTCTTAAGATGA	597
Db	535	CAGGGGATGCTTATATGCTTCTTCACGCGACCATTAACATTCATCTTAAGATGA	594
QY	598	GTTATACGCTAGGAGATGGCTGCTCACAAGCCCTTCATCTGCTCGAAGAAATCTGTC	657
Db	595	GTTATACGCTAGGAGATGCTGCTCACAAGCCCT-----	629
QY	658	AAATTCAGAACCTATGCGCGACAAAATAGCATATAACATTCACAAACAACTGGGTACC	717
Db	630	-----ATTGAATGGGAAAGAGGATCTGCTCCTTCCTTAATGTTAATGAC	681
QY	718	TTCTGTAGCAATTCAGGAATCAATATCTAACACAGCAACAACAGCACTAACAG	777
Db	682	AAATTCAGCTCAACGAATGTCACAGAACACCAATCGATGTCACATCCAACTTATCA	741
QY	778	TCAAGCAAGTCAAGTAATAGCATATATGTCCTTGAACAGCTCTCAAACTGCCCTT	837
Db	742	TCA---AAATCAAGGGGAAAACATTTTCAAGCCTTTTACTGTAATGTATGCTAAACCTT	798
QY	838	GAGTCAACGACATGTAAATCTTGATGCGCTGTCTTTGATCCAGACACAATTCACAGTCG	897
Db	799	ATCAGAAAGCCATGTGAATCTGATGCGCTTATTTTCGACCAGCGCAAAATCACAAATCG	858
QY	898	AACGCTAGAGGTGTGACGTGGCACAGGAGATATTAACGACCTTCATCCCTTACTCTCA	957
Db	859	AACCGCCAGAGGTGTAGCTGTCCCTCATGTGTAACCACTTATACCTTTATCCCTTATGAACA	918
QY	958	AATGTCTGAATTTGGAAGAACGAATCGCTGTATTTATCCCTTTCGTTATCGTTAAACCA	1017

D	b	919	AAATCTGAATTTGGAAAAAGCAATTGCTGTAATTAATCCCTTCGTTATCTGTTCAACCA	978
Q	y	1018	TTGGGTACAGATTCAGAGGCCAGACAGCAAGTCCACACAGCACTCCGGAACTAGTCC	1077
D	b	979	TTGGGTACAGATTCAGAGGCCAGACAGCAAGTCCACATGACTCCGGAACTAGTCC	1038
Q	y	1078	AGGCGCCGAACCTGCACCAATCTTTAAATAGACTCAATTCTTC-----TTT	1125
D	b	1039	AAGTCCCAACCTGCACCAATCTTCACAGCTCCACCAATCAATGATGAGAAATT	1098
Q	y	1126	GGTATGACCTGGTACGAAATTTGGGGAGGATATGATTCGAAAGAAAGGAGATCC	1185
D	b	1099	GGTCAAGAAAGCTGTTCGAAATATAGCGATGTATTTGCTTTGAGGAAATGAGTTTC	1158
Q	y	1246	ATCAAAACAAGAGTGTTCACACACTTTAACTGCTAATAAAGAAATGTTGCTCTCG	1305
D	b	1219	GGCCACACAGAAAGATTATCTATAGCTAGAGCTAAGAAACCTGACCTCCATCTAG	1278
Q	y	1306	TGACCAAGATTTTATGATTAAGCATATATCTGTTAACTAGAGCTCATAAAGCTGTT	1365
D	b	1279	TGATCGAAGATTTTACATATAGGCTTATGACTTACGAGAAATTTACACAGATTTACT	1338
Q	y	1366	TGNAATAAAGAGGTCGTAAATCTATTTCCAGCCTTAGACAAATTAATGAAAGCTTGAA	1429
D	b	1339	TGATATAAAGGTCGACAACTTATTTTGAGCTTTGGATTAACCTGTGTGAAACGACTAA	1398
Q	y	1426	TGATGAATGACATAATTAAGAAAAATTTGGTAGATTTATTTGGCATTCAGACCAAT	1485
D	b	1399	GGATGTCNCAAGTGATTAAGTCAAGTTAGTGANGATATTTGCTCTTATGCTCCGAT	1458
Q	y	1486	TACCATTCAGAGGAGCTTGGCAACCAATTTCAATTTGAGTATCTAGAACGAAGT	1545
D	b	1459	TGCGATCATCGAAGAGTTTATAGAAAACCAATGCGCAAAATTAACCTACAGATGAGGAAT	1518
Q	y	1546	TGCTATTGTCATTAATGACTGATTAAGTATPACACGTCAGATGGTTACATTTTGATGAGCA	1609
D	b	1519	TCAAGTACCAAGTTGGCAGGCAAGTACACAGAGAGGTTTATCTTTTATGATCTCG	1578
Q	y	1606	TGATATTAATCATGATGAGAGAGATCATATGTAAACGCTCATATGAGGCTATGACTG	1665
D	b	1579	TGATATAACAGATGATGAGGGGATGCTGTAATGTAATCCATATGACCATAGGCACCTG	1638
Q	y	1666	GATTGGAAGAGATAGCCTTTCTGATTAAGGAAAAAGTTGCAGCTCAGGCTATCTAANA	1725
D	b	1639	GATTAATAAAGATAGTTTGTCTGAACTGAGAGCGGACGCCAGGCTTATGTTAANA	1698
Q	y	1726	AAAAGTATCTACCTCCACCTCCACAGCCAGATGTTAAAGCAATCCAACTGGAGATAG	1785
D	b	1699	GAAAGGTTTACCCCTCTCCGACAACATCCAGGATTAAGGAAATATCGAGGCAAAAG	1758
Q	y	1786	TGCAGCAGCTATTTACCATGTTGAAAGGGGAAAAACAATTCACCTGTCGACTTCC	1845
D	b	1759	AGCAGAGCATATCTACAAACGCGCTGAAGACGTTAAGAAAGGTCCATGTATGATGCC	1818
Q	y	1846	ATATATGTTGAGCATACAGTTGAGGTTAAAAACGTTAATTTGATTATTTCTCATNAAGA	1905
D	b	1819	TTACAAATCTTCATATTAATCTAGAAAGTCAAAAACGGTAGTTTATATCATATCCATTATGA	1878
Q	y	1906	TCATTAACCATATATTAATTTCTTGTTGTTGATGATCACAATPCCAAAGCTCCAAATGG	1965
D	b	1879	CCATTAACCATATATCAAAATTTAGTGGTTTGACGAAGGCTTTTATGAGGACCTTAAGG	1938
Q	y	1966	CTATACCTTGAAGATTTGTTTGCAGAGATTAAGTCTACGTAGAGAACCTCGAGAGAG	2025
D	b	1939	GTAATCTTGTAGAGATCTTTTGGCAGCTGTCAAGTACTATGTGCAACATCCAAAGCAAG	1998
Q	y	2026	TCCACATCTTAATGATGATGATGGGCAATGCCAGTGACATGTGTTAGCAGAAAGACCA	2085
D	b	1999	TCCGATCTAGATTAATGATTTTGGTAAAGCTACGCGCACTGTTCAAAAGAAACAAAAATGG	2058



```
Db 1099 GGTCAAAAGAGCTGTTGAAAAGTAGCGATGGTTATGCTTTGAGAGAAATGAGATTTC 1158
QY 1186 TCGTTATGCTTTTCGAAAGATTTTACCATGGAATCTTAAAAATCTTGAACCAAGTT 1245
Db 1159 TCGTTATATCCGACGCAAGATCTTTGACAGAAACGACGACGACGATTTGATGCAACT 1218
QY 1246 ATCAAAACAAAGAGAGTGTTCACACACTTAACTGCTAAAAAAGAAATGTGCTCTCG 1305
Db 1219 GGGCAACACGAGAAAGTTTATCTCATAGAGCTAGAGCAAGAACTGACCTCCATCTAG 1278
QY 1306 TGACCAAGAAATTTATATATAAGCATATATCTGTACTGAGGCTCATAAAGCCTTGT 1365
Db 1279 TGATCGAAGATTTTACAAATAGGCTTATGACTTACTAGCAAGATTCACCAAGTTTACT 1338
QY 1366 TGNAAATTAAGGTCGTAAATCTGTTTCCAAAGCTTAGACAATAATATAGAACCTGAA 1425
Db 1339 TGATTAATTAAGGTCGACAACTGATTTTGAGGCTTTGGATTAACCTGTTGGAAAGCTCAA 1398
QY 1426 TGATGAATCGACTAATTAAGAAAAATGTAGATGATTTTGTGCTTCCAGCACCAAT 1485
Db 1399 GGATGTCNCAAGATGATTAAGTCAAGTATGAGANGATATCTTCCCTTCTTACCTCGAT 1458
QY 1486 TACCATCCAGAGGAGCTTGSCAAACCAATTCCTCAATGAGTATACGAGAGAGT 1545
Db 1459 TCGTCATCCAGAAAGCTTAGAAAAACCAATGCGCAATTAACCTACGATGATGAGAT 1518
QY 1546 TCGATTCGCATTAATGCTGATTAAGTATACAGCTCAGATGTTTACATTTTGTGTAACA 1605
Db 1519 TCAAGTATGCGAAGTGGCAGGACAGTACACAGAAAGAGCGCTATATCTTCTGCTCG 1578
QY 1606 TGATATTAATCAGTGAAGAGAGATGATGTATGACGCTCATATGAGCCATAGTACTG 1665
Db 1579 TGATATTAACCATGATGAGGAGGATGCTATGTTATCTCCCATATGACCATACCTAC 1638
QY 1666 GATTGGAAGAGATGCTTTCTGATTAAGAAAAAGTTGCGAGCTCAGCCTTACTATAAGA 1725
Db 1639 GATTAAAAAAGATAGTTGTCTGAAAGCTGAGAGCGCGACGCTTATGCTTAAGA 1698
QY 1726 AAAAGATCTCTTCACCTCCTCAGACGAGAGATGTTAAAGCAATCCAACTGAGAGTAG 1785
Db 1699 GAAAGGTTTAAACCCCTCTTCGACAGACCATCGAGATTCAGAAATCTGAGGCAAAAAG 1758
QY 1786 TGCAGCAGCTATTTCATCAATGCTGTGAAGGGAAGAAAAAGAAATCCACTGCTGACTTCC 1845
Db 1759 AGCAGAACTATCTACAAACCGCGTGAAGCAGCTAAGAGTGCACCTGATGCTATGCC 1818
QY 1846 ATATATGCTTGAGCATACAGTTGAGGTTAAAAACGTTAATTTGATTATTCCTCATAGA 1905
Db 1819 TTCAATCTTCAATATCTACTAGTAGCAAAACGCTAGTTTATCATCTATCTCATTTATGA 1878
QY 1906 TCAATTAACATAATTAATTTGCTTGTGTTGATGATCACACATACAAACCTCCAAATGG 1965
Db 1879 CCAATTACCAATACCAATTAATTTGAGTGTGTTGAGAAAGGCTTATGAGGACACTTAAGG 1938
QY 1966 CTAATACCTTGAAGATTTGTTTGGAGAGATTAAGTACTAGTAGAAGACCTTGAAGAG 2025
Db 1939 GTATACTCTTGAAGATCTTTTGGGAGCTGCAAGTACTATGTGSAATCAATCAAGAGAG 1998
QY 2026 TCCACATCTTAATGATGATGGGCAATGCAAGTGAAGCTGTGTTAGCAAGAAAGCA 2085
Db 1999 TTCGCAATTCAGATTAATGCTTTGTTGTAACGCTAGGACCATGTTCAAAAGAAACAAATATG 2058
QY 2086 CAGTGAAGATCCAAATTAAGAACTTCAAGGAGATGAAGAGCCAGTAGAAGAAACACTGC 2145
Db 2059 TCAAGCTGATACCAATCAAAACGAAAAACCAAGGAGAGAAACCTGACAGAAAAACC 2118
QY 2146 TGAGCCAGAAAGTCCCTCAAGTAGAGA 2171
Db 2119 TGAGGAAGAAACCCCTCGAAGAGAGA 2144
```

RESULT 14

```
AAA65737
ID AAA65737 standard; DNA; 2639 BP.
XX
XX
AC AAA65737;
XX
XX
DT 21-NOV-2000 (first entry)
XX
XX
DE Streptococcus pneumoniae BVH-11-2 gene SEQ ID NO:13.
XX
XX
KW Streptococcus pneumoniae; BVH-3; BVH-11; BVH-28; antigen; vaccine;
KW prophylaxis; therapy; infection; diagnosis; meningitis; bacteraemia;
KW otitis media; pneumonia; immunisation; bactericidal; ds.
XX
XX
OS Streptococcus pneumoniae.
XX
XX
PN NC0200039299-A2.
XX
XX
PD 06-JUL-2000.
XX
XX
PF 20-DEC-1999; 99MO-CA01218.
XX
XX
PR 23-DEC-1998; 98US-0113800.
XX
XX
PA (BIOC-) BIOCHEM PHARMA INC.
XX
XX
PI Hamel J, Brodeur BR, Pineau I, Martin D, Rioux C, Charland N;
DR WPI; 2000-452397/39.
DR P-PDB; AAB12720.
XX
XX
PT Streptococcal antigens useful for vaccinating against e.g. meningitis,
PT otitis media, bacteraemia and/or pneumonia -
XX
XX
PS Example 6; Fig 16; 106pp; English.
XX
XX
CC The present invention describes nucleic acids (i) encoding protein
CC antigens (ii) from Streptococcus pneumoniae. The protein antigens
CC have bactericidal activity. The nucleic acids, encoding the protein
CC antigens, may be used for the recombinant production of the proteins
CC they encode. The protein antigens may then be used as vaccines for the
CC prevention and treatment of Streptococcal infections in mammals
CC (especially humans) which result in, e.g. meningitis, otitis media,
CC bacteraemia and/or pneumonia. The present sequence encodes the
CC S. pneumoniae BVH-11-2 protein antigen.
XX
XX
SQ Sequence 2639 BP; 889 A; 518 C; 567 G; 665 T; 0 other;
XX
XX
Query Match 41.1%; Score 980.8; DB 21; Length 2639;
Best Local Similarity 67.8%; Pred. No. 2.5e-236;
Matches 1473; Conservative 0; Mismatches 638; Indels 60; Gaps 5;
QY 1 TTCTTACGATTTGGGACTGATATCAAGCTAGAACGTTTAAGAAAA--TTATGCTGTTTC 57
Db 173 TTCTTATGAACCTTGCTGTGTCACCAAGCTGTGCAAGTTTAAGAAAGTCTATCGAGTTTC 232
QY 58 CTAATATGATGAGAAACAAAGCAGACGCAAAAAAGAGATTTGACCTGATGAGGTTAG 117
Db 233 TTATATGATGATGATGATGAGCTGTCTCAAAAGCGCAAGAAATTTGACACCAATGAGCAG 292
QY 118 CAAGCGTGAAGAAATCAATGCTGAGCAATGCTCATCAAGATTAACAGCAAGGCTATGT 177
Db 293 TAAGAGAGAGGAGATCAACGCCGAAACAAATTTGATATCAAGATTAAGGATCAAGGTTATGT 352
QY 178 CACTTCATGAGGAGCAGCATATCTATTATTAATGTAAGTGTCTTATGACGTATCAT 237
Db 353 GACCTTCATGAGAGACCATTAATCACTATTAAGGAGAGGCTCTTATGATGCAACAT 412
QY 238 CAGTGAAGAAATTAATCAATGAAGATCAAACTAATAGCTAAAGATGAGATATTGTTAA 297
Db 413 CAGTGAAGAACTCTCATGAAGATCGGAATTTATCAAGTGAAGATTCACACATGTGCAA 472
QY 298 TGAAGTCAAGGTTGAGATATTTATCAAGTAGATGAAAAATCACTATGTTTACCTTAAGA 357
```

473 TGAATCAAGGGTGGCTATGTGATTAAGTAGACGAAAAAATCTATGTATACCTTTAAAGA 532  
QY 358 TGGTCCCGACGCGGATACGTCCTGTACAAAAGAGAAATCAATGCACAAAACAGAGCA 417  
DB 533 TGGGGCCCATGCGGACAAATTTGCGACAAAAGAGATTTAAAGCTGAGAGAGAGACA 592  
QY 418 TACTCAACATCGTGAAAGGTGAACTCCAGAAACGATGGTGTGTCCTTGGCACGTTTC 477  
DB 593 CACTCATTAATCAT- - - - -AACTCAAGACGATTAATGCTGTCTGCACGCAAGC 643  
QY 478 GCAAGAGCGCTACTACTACAGATGAGTGTATATCTTTAATGCTTGTGATATCATAGAGA 537  
DB 644 CCAAGGACGTTATACACGAGATGAGGTATATCTTCAATGCACTGTATATCATGTAGGA 703  
QY 538 TACTGTGATGCTTATATCTCTCTCATGAGATCATTTACCATTTACCTTAAGATGA 597  
DB 704 CACGGGTGATGCTTATATCTCTCTCATGAGGACCATTTACCATTTACCTTAAGATGA 763  
QY 598 GTTATCAGCTAGGAGTGTGCTGCTGCAGAAAGCCTTCTATCTGTGTGAGGAAAATCTGTC 657  
DB 764 GTTATCAGCTAGGAGTGTGCTGCTGCAGAAAGCCT- - - - - 798  
QY 658 AATTTCAAGAACTATTCGCGGACAAATAGCGATTAACATTTCAAGAACAACTGGTACC 717  
DB 799 - - - - -AATGGAATGGAGAGAGAGGATCTGCTTCTTCAAGTTCTAGTTATTAATGC 850  
QY 718 TTCTGTAGCAATTCAGGAATCTCAATTAATAACACAGAACAGACAGCAACTATACAG 777  
DB 851 AATTCAGTTTCAACCAAGATTTGTCAGGAACCAATCTACGTCTCATCTCCAACTTATCA 910  
QY 778 TCAAGCAGTCAAAAGTATGATGATGCTCTTGAAGAACGCTTACAAATGCTCTT 837  
DB 911 TCA- - - - -AAATCAAGGGGAAAAATTCACACCTTTTACGTGAATGTGATCTTAACCTT 967  
QY 838 GAGTCAACAGTGTAGATCTGTATGAGCTTGTCTTTGATGACGACAAATTCACAGTGC 897  
DB 968 ATAGAAAGCGCATGTAGATCTGTATGAGCTTATTTTGCACGACGCAATTCACAGTGC 1027  
QY 898 AACAGCTAGAGGTGTGCACTGTCACACGAGATCATTTACACTTTCCTTACTCTCA 957  
DB 1028 AACCGCAGAGGTGTAGCTGTCTCCCTATGTTAACCATTTACACTTATGCCCTTATGAGA 1087  
QY 958 AATGCTGATTTGGAAGAGCATTCGCTGATATTTCCCTTCTGTTATGTTCAACCA 1017  
DB 1088 AATGCTGATTTGGAAGAGCATTCGCTGATATTTCCCTTCTGTTATGTTCAACCA 1147  
QY 1018 TTGGGTACCGATTTCAAGGCGAGAACACCAAGCTCCGACGACGACGACGACGACGAC 1077  
DB 1148 TTGGGTACCGATTTCAAGGCGAGAACACCAAGCTCCGACGACGACGACGACGACGAC 1207  
QY 1078 AGCCCGCAACCTGCAACCAATCTTAATATGACTCAAAATTTCTT- - - - -TTT 1125  
DB 1208 AATCTGCAACCTGCAACCAATCTTCAACGAGTCCCAAGCAATTCATGATGAGAAAT 1267  
QY 1126 GGTATGCTAGCTGTACGAAAGTTGGGAGAGATATGTTGAAAGAAAAGGCAATCTC 1185  
DB 1268 GGTCAAGAAAGCTGTTCGAAAGTAGGCGATGTTATGTTGAGAGAGATGAGATTTC 1327  
QY 1186 TCGTTATGCTTTGGCAAGATTTACCATCTGAACTGTAAAAATCTTGAAGCAAGTT 1245  
DB 1328 TCGTTATGCTTTGGCAAGATTTTTCAGAGAACGACGAGGAGATGATGCAAACT 1387  
QY 1246 ATCAAAACAGAGAGTGTTCACACACTTAACTGCTAAAAAGAAAATGTCTCTCTCG 1305  
DB 1388 GGCACAGCAGGAAGTTATCTATAGCTAGAGACTAGAAAACGCTACCTCATCTAG 1447  
QY 1306 TGACCAAGAAATTTATGATTAAGCATATATCTGTTAATGAGGCTCATTAAGCCCTGTT 1365  
DB 1448 TGATCGAATTTTCAATTAAGGCTTATGACTTACGAAAGTTCACACAGATTTACT 1507  
QY 1366 TGAATAAGAGGCTGATATTTCTGATTTCCAGGCTTAGACAAATTTATGAAAGCTTGA 1425  
DB 1508 TGATTAATAAGGTCGACAAAGTTGATTTTGAGTTTGATTAACCTGTGTGAACGACTCA 1567

QY 1426 TGATGAATGACTAATTAAGAAAAATTTGATAGATTTATTTGGCATTCCTTACACCAAT 1485  
DB 1568 GATGTCTCAAGTGATTAATTAAGTATAGTATGATATTTCTGCTCTTACTCTCGAT 1627  
QY 1486 TACCATTCAGACGACTTGGCAACCAATTTCTCAATTTGAATATAGTGAAGCAAGT 1545  
DB 1628 TCGTCACTGCAAGAGTTTAGGAAACCAATTCGCAATTTACCTATGATGATGAT 1687  
QY 1546 TCGTATTTGCTCAATTTAGCTGATAGTATTAACAGCTGATGATGATTTGATGAAC 1605  
DB 1688 TCAAGTATGCAAGTTGACAGGCAATGACACAGACAGAGGTTATATCTTATCTCTG 1747  
QY 1606 TGATATTAATCAGTATGAAAGAGATGATATTAAGCCTCATATGAGGCTATGACTG 1665  
DB 1748 TGATATTAACAGTATGAGGAGGATGCTATGATATCCCATATGACCATGAGCCACTG 1807  
QY 1666 GATTGCAAAAGATAGCCTTTCTGATTAAGAAAAGTTGACGCTCAAGCCTATATCTAAGA 1725  
DB 1808 GATTAATAAAGATAGTTTGTCTGAAGCTGAGAGAGGCGCAGCCAGGCTTATCTAAGA 1867  
QY 1726 AAAAGTATCTACCTCCATCTCCAGCGAGATGTTAAAGCAAAATCCAACTGGAGATAG 1785  
DB 1868 GAAAGGTTGACCCCTCTCTGACAGACACAGATTCGAAATATCTGAGGCAAAAG 1927  
QY 1786 TGCAGCAGCTATTTACAAATCGTGTGAAGGGAAGAAAACGAATTCACCTGTTGACTTCC 1845  
DB 1928 AGCAGAAAGCTATCTACACGCGCTGAAGACGACTAAGAAAGTCCACTTATGTATGCC 1987  
QY 1846 ATATATGTTGAGCATACAGTGTAGAGTTAAAGCGTAATTTGATTTATCTCATTAAGA 1905  
DB 1988 TTCAATCTTCAATATATCTGTAGAAAGTCAAAAACGGTAGTTTATCATCTCATTAATGA 2047  
QY 1906 TCATTTACATTAATTAATTTGCTTGTGATGATTCACATATACATAAGCTCCAAATG 1965  
DB 2048 CCATTTACATTAATTAATTTGATGATTTGACGAAGGCTTTATGAGGACACTTAAGG 2107  
QY 1966 CTATACCTTGGAGATTTGTTGCGACGATTAAGTACTACGTAGAAACACCTGACGAG 2025  
DB 2108 GTATAGCTTGAAGATCTTTTGGCGACTGTCAGTCACTATGTGAAACATCCAAACGAG 2167  
QY 2026 TCCACATTTCAATGATGATGAGGCAATGCGACTGAGCATGTGT7AGCAAGAAAGCA 2085  
DB 2168 TCCGATTTCAATGATGATGATTTGTTGTAAGCTGATGACATGTCTGTAATAATTAAGCAG 2227  
QY 2086 CAGTGAAGATCCAAATTAAGACTTCAAGGAGTGAAGAGCCGATGAGGAACACACTG 2145  
DB 2228 CCAAGATTAATTAACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2287  
QY 2146 TGAAGCCAGAG 2156  
DB 2288 TGAATCTGATG 2298

RESULT 15  
ABK15104  
ID ABK15104 standard; DNA; 2639 BP.  
XX  
AC ABK15104;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE DNA encoding Streptococcus pneumoniae BVH-11-2.  
XX  
KW BVH-3; BVH-11; vaccine; meningitis; otitis media; bacteremia;  
KW pneumonia; streptococcal bacterial infection; gene; ds; BVH-11-2.  
OS Streptococcus pneumoniae.  
XX  
FH key location/Qualifiers  
FT CDS 114..2630  
FT /tag= a  
FT /product= "BVH-11-2"



Db 1508 TGAATAAAGGTCGACAGTGTGATTTTGAGGTTTGATTAACCTGTTGGAAAGACTCAA 1567  
QY 1426 TGATGATTCGACATTAATAAGAAAAATGGTAGATGATTTATGGCATTCCTAGCACAAT 1485  
Db 1568 GGATGTCTCAAGTGAATAGTCAGTAGTGATGATATTTGCTTGTAGCTCCGAT 1627  
QY 1486 TACCATCCAGAGGACTTGGCAACCAATTCYCAATTGAGTATACTGAAGACGAGT 1545  
Db 1628 TCGTCAATCCAGAGCTTGGAAAAACCAATGCCCAATTACCTACACTGATGATGAGAT 1687  
QY 1546 TCGATTCCTCAATTAAGTATACAGCTCAGATGTTTACATTTTGTGATGACA 1605  
Db 1688 TCAAGTAGCCAGTGGCAGGACGACACAGAAAGACGTTATATCTTGTGATCCTCG 1747  
QY 1606 TGATATATCGATGATAGAGAGATGATATGTAACGCTCATATGGGCCATAGTCACTG 1665  
Db 1748 TGATATACCAAGTGAAGGAGGAGTGCCTATGTAAGTCAATATGACCAATGACCACTG 1807  
QY 1666 GATTGAAAAAGATPAGCCTTCTGTAAGAAAAAGTTGCAGCTCAAGCTATATAAGA 1725  
Db 1808 GATTAAAAAGATGTTGTGCTGAAGCTGAGAGGCGGACGCCAGGCTTATGCTAAGA 1867  
QY 1726 AAAAGTATCCTAAGCTCCATCTCCAGACGAGATGTTAAAGCAAAATCCAAGTGAAGTAG 1785  
Db 1868 GAAAGGTTTGAACCTCTCCAGACGACCAAGATTCAGAAATACTGAGGCAAAAGG 1927  
QY 1786 TGCAGCAGCTATTTACATGATGTGAAAGGGGAAAAAGAAATCCACTCGTTCGACTCC 1845  
Db 1928 AGCAGAAAGCTATCTACACGCGTGAAGACGACTAAAGAGGTCGACCTGATGATGCC 1987  
QY 1846 ATATATGTTGACATPACAGTGTAGGTTAAAGCGTAATTTGATATTCCTCATAGA 1905  
Db 1988 TTACAAATCTCAATATATCTAGAACGCAAAAGCGTAGTTAATCATACCTCATTTGA 2047  
QY 1906 TCATTTACCAATATTAATTTGCTGTTGATGATCAGACATACAAGCTCCAAATGG 1965  
Db 2048 CCATTACCATACATCAAAATTTGAGTGTGACGAAGGCTTTATGAGGCACCTAAGGG 2107  
QY 1966 CTATACCTTGGAGATTTGTTGGGAGATTAAGTACTAGTAGAACACCTGACGAAG 2025  
Db 2108 GTATAGCTTGAAGATTTTGGGAGCTGCAAGTACTATGTCGACATCCAAACGAAAG 2167  
QY 2026 TCCACATTTCAATGATGATGGGCAATGCCAGTGAAGATGTTAGCAAGAAAGACA 2085  
Db 2168 TCCGATTTCAAGTAATGTTTGTGAACGCTAGTACCATGTTGTAATAAATAAGGACA 2227  
QY 2086 CAGTGAAGATCCAAATAGACTTCAAGCGGATGAAGAGCCAGTAGAGGAAACCTGC 2145  
Db 2228 CCAAGATAGTAACTGATGAAGTAAGGAACATGATGAAGTGAAGGCAACTCACCC 2287  
QY 2146 TGAGCCGAGAG 2156  
Db 2288 TGAATCTGATG 2298

Search completed: November 17, 2002, 01:50:41  
Job time : 373 secs